ACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGGATCCTCTAGAGATCCCTCGACCTCGA CCCACGCGTCCGGGCCGGAGCACGCCGCAGGACCTGGAGCTCCGGCTGTCTCCCG  $\texttt{CAGCGCTACCCGCC} \underline{\textbf{ATG}} \texttt{CGCCTGCCGCGCCGCGCCGCGCTGC} \underline{\textbf{ATG}} \texttt{CTCCTGCCGCTTCTGCTG}$ CTGCTGCCGCCCGGGGGGCCCAAGAAGCCGACGCCCTLCCACCGGTGCCGGGGGCT GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAÄCTTTGGCGGCGGGAACA CGGCTTGGGAGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTCGCCTGCTGGAGATC CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGTGCT GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGT GGCTGCACAGGGAAGGCCCAGGAAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA CGGACAGTGTGCAGATGTGGACGAGTGCTCACTAGCAGAAAAACCTGTGTGAGGAAAAACG GAAGATGCCTGTGTGCCGCCGGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT  ${\tt GCCCTCCCGCGAAGACCTG}$  ${\tt CGTGGAAAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGGGGGAGAGGCTGC}$ CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG AGGAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCGACCTGCAGAAGC TTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA TCACAAATTCACAAATAAAGCATTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAACTC ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCGGCGCAGCACCATGGCCTGAAAT AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG ATCTCAATTAGTCAGCAACCCAGTTTT



# 

### FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop ><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEEKTLSKYESSEIRL LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGGSQRPCSG NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESCKTCSGLTNRDCGECEVGWVLDE GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT CVRKNENCYNTPGSYVCVCPDGFEETEDACVPPAEAEATEGESPTQLPSREDL

#### Signal peptide:

amino acids 1-24

N-glycosylation sites.

amino acids 190-194 and 251-255

Glycosaminoglycan attachment sites.

amino acids 149-153 and 155-159

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 26-30

Casein kinase II phosphorylation sites.

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343 and 349-353

Tyrosine kinase phosphorylation site.

amino acids 303-310

N-myristoylation sites.

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and 313-319

Aspartic acid and asparagine hydroxylation site.

amino acids 308-320

EGF-like domain cysteine pattern signature.

amino acids 166-178

Leucine zipper pattern.

amino acids 94-116

CAGGTCCAACTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCCCAGCCGTCTAAACGGGAACA GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCCTGAGCAGCATGGCCCGGAGGAGCGCCTTC CGGGCCGCCGCAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA TAGGATTTGAAGAAGATATCCTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGAT TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCCTGTCAATATCCATTCCATGAATTTTAC CTGGCAAGCTGCAGGCAGGCAGAATACTTCTATGAATTCCTGTCCTTGCGCTCCCTGGATA AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA TCAGTTGTTCAAGTTGGTTTCCCATGTCTTGGAAAACAGGATGGGGTGGCAGCATTTGAAGT GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT TCTTTAAAACATGTCAACAAGCTGAGTGCCCAGGCGGGTGCCGAAATGGAGGCTTTTGTAAT GAAAGACGCATCTGCGAGTGTCCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG TACCCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCCACCTG GATTCTATGGAGTGAACTGTGACAAAGCAAACTGCTCAACCACCTGCTTTAATGGAGGGACC TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG CAAATGCCCACAACCCTGTCGAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT GGAACCTGCCATGAACCCAACAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA TAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCCCAGCTCAGGCAGC ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG **TGA**ACTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTCA TGTGTTGAATGTTCAAATAATGTTCATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT TCATTATAAATCACTGAGCTGATATTTACTCTTCCTTTTAAGTTTTCTAAGTACGTCTGTAG CATGATGGTATAGATTTTCTTGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA TCAGGTTAAAATTTTCAGTGTGTAGTTGGCAGATATTTTCAAAATTACAATGCATTTATGGT GTCTGGGGGCAGGGAACATCAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAAT TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT GTTTGTTACATTTTTAAAAATTGCTCTTAATTTTTAAACTCTCAATACAATATATTTTGACC AAACAATATAATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCAT TGGCTTGAAGCAATATATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT AAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGC CGCCATGGCCCAACTTGTTTATTGCAGCTTATAATG

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094
><subunit 1 of 1, 379 aa, 0 stop
><MW: 41528, pI: 7.97, NX(S/T): 2</pre>

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFLSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFHGPHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAEERRDP
PESNYIW

Signal peptide:

amino acids 1-28

N-glycosylation site.

amino acids 88-92, 245-249

Casein kinase II phosphorylation site.

amino acids 319-323

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

CGGACGCGTGGGCGTCCGCAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG TGCTGTTGCTGCCGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCGTGGGCGCCGG GAGCTAGCACCGGGTCTGCACCTGCGGGGCCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA GCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCCCTGACTTCTGGGACTTC TGCCTCGGCGTGCCACCCCCTTTTCCCCCCGATCCAAGGATGTATGCATGGAGGTCGTATCTA TCCAGTCTTGGGAACGTACTGGGACAACTGTAACCGTTGCACCTGCCAGGAGAACAGGCAGT  ${ t GAACCACAGCGCCTTCTGGGGCATGACCCTGGA}{ t { t CGC}}{ t { t CGCTACCGCCTGGGCACCA}$ TCCGCCCATCTTCCTCGGTCATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGA CCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT CAATCCATTCTCTGGGACACATGACGCCTGTCCTGTCGCCCCAGAACCTGCTGTCTTGTGAC ACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCG CCGAGGGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCC CTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGCC CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCT CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA TGGAGGTGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGC CTTGGGAGGCCAGAGATACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC CAGCCTGGGGCGAGAGGGGCCACTTCCGCATCGTGCGCGCGTCAATGAGTGCGACATCGAG AGCTTCGTGCTGGGCCTTGGGGCCGCGTGGGCATGGAGGACATGGGTCATCACTGAGGCTG CGGGCACCACGCGGGGTCCGGCCTGGGATCCAGGCTAAGGGCCCGGAGGGCCCCAATG GGGCGGTGACCCCAGCCTCGCCCGACAGAGCCCGGGGCGCGAGGCGGCGCCAGGGCGCTAAT CCCGGCGCGGGTTCCGCTGACGCAGCGCCCCGCCTGGGAGCCGCGGGCAGGCGAGACTGGCG GAGCCCCCAGACCTCCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAG ATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC  ${ t TTGCCCAGGTTGGAGTGGCCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA}$  ${f AGTGACCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACCTGGC}$ TAATTTTTGTATTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCCAGGCTGGTTTCGAACT CCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCC TAAAACCAAAGTATTGATAAAAAAAAA

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223</pre>

><subunit 1 of 1, 164 aa, 1 stop

><MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQDLCCRGRADDC ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPPFPPIQGCMHGGRIYPVLGTYWDNCNR CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

N-glycosylation site.

amino acids 78-82, 161-165

Casein kinase II phosphorylation site.

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300, 411-415

N-myristoylation site.

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230, 269-275, 378-384, 442-448

Amidation site.

amino acids 26-30, 318-322

Eukaryotic thiol (cysteine) proteases histidine active site. amino acids 398-409

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTCGTTGTCTCAAATCCA
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGG
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC
CCCCTGTATGATGCACAGCCGAGCCATGGGTCGGGGCAAGCGCCAGGCCACTGC
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCTCGGCTC
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG
TGCATGAGGACTTCTTCCTATACAAGGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG
AGGCCAGAGAGAATACCGCCGGCATGGGACCCCACTCAG

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG  ${\tt CAGCTTGAACTGCGTGG} \underline{{\tt ATG}} {\tt ACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT}$ GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGGACCCCGGCCAGCTATAGGCTCTGGGGGG CCCCGCTGCAGCCCACACTGGGTGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG CTGCACCCCTGTCCCCACCCTGACCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC CATGGCCCAGCATTCTCCACCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT CCCTGCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA  $\tt GGGGACAGGCACTCAGGAGGGCCCAG{\color{red}{\textbf{TAA}}} AGGCTGAGATGAAGTGGACTGAGTAGAACTGGA$ GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA  ${\tt GGGGCCAGGCCTCACATTCGTGGGGCCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT}$ AATAAACACCTGTTGGATAAGCCAAAAAA

MTHRTTTWARRTSRAVTPTCATPAGPMPCSRLPPSLRCSLHSACCSGDPASYRLWGAPLQPT LGVVPQASVPLLTDLAQWEPVLVPEAHPNASLTMYVCTPVPHPDPPMALSRTPTRQISSSDT DPPADGPSNPLCCCFHGPAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVVSPAPSRGQALRRAQ

Signal peptide:

amino acids 1-47

N-glycosylation site.

amino acids 31-35, 74-78, 84-88

Casein kinase II phosphorylation site.

amino acids 22-26, 76-80

N-myristoylation site.

amino acids 56-60

Amidation site.

amino acids 70-74

 ${\tt CCCACGCGTCCGAACCTCTCCAGCG}$ GCTTACAGCTGCTGATTCTCTGCTGTCAAACTCAGTACGTGAGGGACCAGGGCGCCATGACC GACCAGCTGAGCAGGCGGCAGATCCGCGAGTACCAACTCTACAGCAGGACCAGTGGCAAGCA CGTGCAGGTCACCGGGCGTCGCATCTCCGCCACCGCCGAGGACGGCAACAAGTTTGCCAAGC TCATAGTGGAGACGGACACGTTTGGCAGCCGGGTTCGCATCAAAGGGGGCTGAGAGTGAGAAG TACATCTGTATGAACAAGAGGGGCAAGCTCATCGGGAAGCCCAGCGGGAAGAGCCAAAGACTG CGTGTTCACGGAGATCGTGCTGGAGAACAACTATACGGCCTTCCAGAACGCCCGGCACGAGG GCTGGTTCATGGCCTTCACGCGGCAGGGGCGGCCCCGCCAGGCTTCCCGCAGCCGCCAGAAC CAGCGCGAGGCCCACTTCATCAAGCGCCTCTACCAAGGCCAGCTGCCCTTCCCCAACCACGC  ${\tt GGCGGCCCCAGCCCTCACG}$ GAGGGAGGACCCTGAGGGCCGCGAAGCATCCGAGCCCCAGCTGGGAAGGGGCAGGCCGGTG CCCCAGGGGCGGCTGGCACAGTGCCCCCTTCCCGGACGGGTGGCAGGCCCTGGAGAGGAACT GAGTGTCACCCTGATCTCAGGCCACCAGCCTCTGCCGGCCTCCCAGCCGGGCTCCTGAAGCC CGCTGAAAGGTCAGCGACTGAAGGCCTTGCAGACAACCGTCTGGAGGTGGCTGTCCTCAAAA TCTGCTTCTCGGATCTCCCTCAGTCTGCCCCCAGCCCCCAAACTCCTCCTGGCTAGACTGTA AGGGTTGTCCACTCCTCACATTCCACGACCCAGGCCTGCACCCCACCCCCAACTCCCAGCCC CGGAATAAAACCATTTTCCTGC

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLIGKPSGKSKDCVFTEIVLE NNYTAFQNARHEGWFMAFTRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEF VGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCCTGTTGAATTTTTTTGCACATGGAG GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC CTAGTGTGCCGCTGCGACAGGAACTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC AATATTCAGACCATTTCACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT GGATGACAACTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGGAGGCTATTAGCC TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGAC TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCATATCCGACATGGCCTTCCAGAA TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC CACCCTCCTCCCGATCTCCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC ACTGCTCGGAATAACCCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCCTGTCCCACCACGACCCCCGGCCTG CCTCTCTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTAT TCCAAACCCTAGCAGAAGCTACACGCCTCCAACTCCTACCACATCGAAACTTCCCACGATTC CTGACTGGGATGGCAGAAAGAGTGACCCCACCTATTTCTGAACGGATCCAGCTCTCTATC CAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGGCATCGTTCAGGAGCGCATAG TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA  $\tt CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGGCTTGATCGGGGGGCGCGGTGATATTT$ GTGCTGGTGGTCTTGCTCAGCGTCTTTTGCTGGCATATGCACAAAAAGGGGCGCTACACCTC CCAGAAGTGGAAATACAACCGGGGCCGGCGGAAAGATGATTATTGCGAGGCAGCCACCAAGA AGGACAACTCCATCCTGGAGATGACAGAAACCAGTTTTCAGATCGTCTCCTTAAATAACGAT CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC ACTGCCATACG<u>TGA</u>CAGCCAGAGGCCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAAGTG CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTGTAACTCTTTGCTTTTTAAATCTT

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI SRAALAQLLKLEELHLDDNSISTVGVEDGAFREAISLKLLFLSKNHLSSVPVGLPVDLQELR VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN PWFCDCSIKWVTEWLKYIPSSLNVRGFMCQGPEQVRGMAVRELNMNLLSCPTTTPGLPLFTP APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

E.

:I

II.

4

as ##

\_ ma

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,

522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

ATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCCCCTGCC  ${\tt TCGCTTCCCAGGCGCGGCGGCTGCAGCCTTGCCCCTCTTGCTCGCCTTGAAA} \underline{{\tt ATG}} {\tt GAAAAGATGCTCGCAGGCT}$  $\overline{\mathtt{GCTTTCTGCTGATCCTCGGACAGATCGTCCTCCTCCTGCCGAGGCCAGGGAGGTCACGTGGGAGGTCCATCT}$  ${\tt TGGTTTTCATCATGACAGCTCTCGCAGTGTCAACACCCCATGACTATGCAAAGGTCAAGGAGTTCATCGTGGACA}$ TCTTGCAATTCTTGGACATTGGTCCTGATGTCACCCGAGTGGGCCTGCTCCAATATGGCAGCACTGTCAAGAATG AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGCTGTCAAGAGGATGCGGCATCTGTCCACGG GCACCATGACTGGGCTGGCCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCCCGGCCCCTGA GGGAGAATGTGCCACGGGTCATAATGATCGTGACAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA  ${\tt AGGCACGGGACACGGGCATCCTAATCTTTGCCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG}$  ${\tt GGAGTGAGCCCCATGAGGACCATGTCTTCCTTGTGGCCAATTTCAGCCAGATTGAGACGCTGACCTCCGTGTTCC}$ AGAAGAAGTTGTGCACGGCCCACATGTGCAGCACCCTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTG GCTCATACGTCTGCAGGTGCAAACAAGGCTACATTCTCAACTCGGATCAGACGTCTGCAGAATCCAGGATCTGT GTGCCATGGAGGACCACAACTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGGCTGTGGACTACTGTGCCTCAGAAAACCACGGATGTGAAC ATGAGTGTGTAAATGCTGATGGCTCCTACCTTTGCCAGTGCCATGAAGGATTTGCTCTTAACCCCAGATGAAAAA CGTGCACAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCGTCAACATGGAGGAGAGCT ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAAACCTGCAGCCGAGTGGACCACTGTGCAC  ${\tt AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCCTTCGTCTGCCAGTGCTCAGAAGGCTTCC}$  ${\tt TCATCAACGAGGACCTCAAGACCTGCTCCCGGGTGGATTACTGCCTGAGTGACCATGGTTGTGAATACTCCT}$ GTGTCAACATGGACAGATCCTTTGCCTGTCAGTGTCCTGAGGGACACGTGCTCCGCAGCGATGGGAAGACGTGTG  ${\tt CAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAACATTCGTGTGAAGCAGTGAAGATTCGTTTGTGT}$  ${\tt GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTATAG}$  ${\tt ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACACGTGCGAGTGCTTGGAGGGATTCCGGCTCG}$ CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAAATCAACCCACCATGGCTGCGAACACATTTGTGTTA  ${\tt ATAATGGGAATTCCTACATCTGCAAATGCTCAGAGGGATTTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT}$  $\tt GCACTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAGAATTTTGAGGTCG$  ${\tt TGAAGCAGTTTGTCACTGGAATTATAGATTCCTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGT}$ ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCAGCCAAAGACATGAAAAAAAGCCGTGGCCC ACATGAAATACATGGGAAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTTACCCAAG ATGAGATAAGTGAAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG CAGGGGAACTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTCACCATAAATATCCAAGACCTACTTT  ${\tt CCTGTTCTAATTTTGCAGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT}$ CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACACGATCAATGCAAATGTGAAAACCTTATAATGT TCCAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC TGGAAAATCGCCTGAGATACAGA<u>TGA</u>AGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGGATTACAAT  ${ t GAACGCAGTGCAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAACAATCAGTACTGA}$ GAAACCTGGTTTGCCACAGAACAAAGACAAGAAGTATACACTAACTTGTATAAATTTATCTAGGAAAAAATCCT  ${\tt AACTTGCTTCTGCCTCATCCTGCCTTAGTGTGCAATCTCATTTGACTATACGATAAAGTTTGCACAGTCTTACTT}$ CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAAAATTCACCACTTCAG

MEKMLAGCFLLILGQIVLLPAEARERSRGRSISRGRHARTHPQTALLESSCENKRADLVFII DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV KRMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKPGC EHECVNMEESYYCRCHRGYTLDPNGKTCSRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI NEDLKTCSRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTCAKLDSCALGDHGCE HSCVSSEDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA EDGKRCRRKDVCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLQYSTQVHTEFTLRNFNSAKDMKKAVA HMKYMGKGSMTGLALKHMFERSFTQGEGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN GITMYAVGVGKAIEEELQEIASEPTNKHLFYAEDFSTMDEISEKLKKGICEALEDSDGRQDS PAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQRMEALENRLRYR

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 221-225

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 115-119, 606-610, 892-896

Casein kinase II phosphorylation site.

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247, 401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784, 781-785, 819-823, 866-870

N-myristoylation site.

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500, 639-645, 690-696, 752-758, 792-798

Amidation site.

amino acids 314-318, 560-564, 601-605

Aspartic acid and asparagine hydroxylation site.

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464, 540-546, 581-587

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGCAGCCTCG  ${\tt C\underline{ATG}} {\tt ATTTCCCTCCCGGGGCCCCTGGTGACCAACTTGCTGCGGTTTTTTGTTCCTGGGGCTGA}$ GTGCCCTCGCGCCCCCCCGCGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGCACGGGGAGGTGTCTTC AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT GTGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCCAC AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT TTCCTCTGTCACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCCCACCTCTC TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGA AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCTGACTCCTTATGAAGCCAGCTG CTGAAATTAGCTACTCACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCCAGGC CCCCTTGATCTGTACCCCACCCCTATCTAACACCCCCTTGGCTCCCACTCCAGCTCCCTGT  ${ t ATTGATATAACCTGTCAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC}$ TGTTTGTATGAAAA

# The transfer of the transfer o

# FIGURE 17

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSCQSPRSKPAVQYQ WDRQLPSFQTFFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLPWPKSSDTISKNGTL SSVTSARALRPPHGPPRPGALTPTPSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSR MGAVPVMVPAQSQAGSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262, 262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 164-175

# 

# FIGURE 18

CGCCACCACTGCGGCCACCGCCAATGAAACGCCTCCCGCTCCTAGTGGTTTTTTTCCACTTTG TTGAATTGTTCCTATACTCAAAATTGCACCAAGACACCTTGTCTCCCAAATGCAAAATGTGA AATACGCAATGGAATTGAAGCCTGCTATTGCAACATGGGATTTTCAGGAAATGGTGTCACAA TTTGTGAAGATGATAATGAATGTGGAAATTTAACTCAGTCCTGTGGCGAAAATGCTAATTGC ACTAACACAGAAGGAAGTTATTATTGTATGTGTGTACCTGGCTTCAGATCCAGCAGTAACCA AGACAGGTTTATCACTAATGATGGAACCGTCTGTATAGAAAATGTGAATGCAAACTGCCATT TAGATAATGTCTGTATAGCTGCAAATATTAATAAAACTTTAACAAAAATCAGATCCATAAAA GAACCTGTGGCTTTGCTACAAGAAGTCTATAGAAATTCTGTGACAGATCTTTCACCAACAGA TATAATTACATATAGAAATATTAGCTGAATCATCTTCATTACTAGGTTACAAGAACAACA CTATCTCAGCCAAGGACACCCTTTCTAACTCAACTCTTACTGAATTTGTAAAAACCGTGAAT AATTTTGTTCAAAGGGATACATTTGTAGTTTGGGACAAGTTATCTGTGAATCATAGGAGAAC ACATCTTACAAAACTCATGCACACTGTTGAACAAGCTACTTTAAGGATATCCCAGAGCTTCC AAAAGACCACAGAGTTTGATACAAATTCAACGGATATAGCTCTCAAAGTTTTCTTTTTTGAT TCCAAAGAGAAAAGCTGCATATGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATA AGAGTATTGGTCCTTTGCTTTCATCATCTGACAACTTCTTATTGAAACCTCAAAATTATGAT AATTCTGAAGAGGAGAAAGAGTCATATCTTCAGTAATTTCAGTCTCAATGAGCTCAAACCC ACCCACATTATATGAACTTGAAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATA GGTATAGGAGTCTATGTGCATTTTGGAATTACTCACCTGATACCATGAATGGCAGCTGGTCT TCAGAGGGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCT GACACATTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTC TTACAAGGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTT ACCTTCTGGTTCTTCAGTGAAATTCAAAGCACCAGGACAACAATTCACAAAAATCTTTGCTG TAGCCTATTTCTTGCTGAACTTGTTTTTCTTGTTGGGATCAATACAAATACTAATAAGCTCT TCTGTTCAATCATTGCCGGACTGCTACACTACTTCTTTTTAGCTGCTTTTTGCATGGATGTGC ATTGAAGGCATACATCTCTATCTCATTGTTGTGGGTGTCATCTACAACAAGGGATTTTTGCA TAGGATACAGATATTATGGCACAACCAAAGTATGTTGGCTTAGCACCGAAAACAACTTTATT TGGAGTTTTATAGGACCAGCATGCCTAATCATTCTTGTTAATCTCTTGGCTTTTTGGAGTCAT TAAGGTCTTGTGCAAGAGGAGCCCTCGCTCTTCTGTTCCTTCTCGGCACCACCTGGATCTTT GGGGTTCTCCATGTTGTGCACGCATCAGTGGTTACAGCTTACCTCTTCACAGTCAGCAATGC TTTCCAGGGGATGTTCATTTTTTTATTCCTGTGTGTTTTATCTAGAAAGATTCAAGAAGAAT ATTACAGATTGTTCAAAAATGTCCCCTGTTGTTTTGGATGTTTAAGG<u>T**AA</u>ACATAGAGAATG</u>** GTGGATAATTACAACTGCACAAAAATAAAAATTCCAAGCTGTGGATGACCAATGTATAAAAA TGACTCATCAAATTATCCAATTATTAACTACTAGACAAAAAGTATTTTAAATCAGTTTTTCT GTTTATGCTATAGGAACTGTAGATAATAAGGTAAAATTATGTATCATATAGATATACTATGT TTTTCTATGTGAAATAGTTCTGTCAAAAATAGTATTGCAGATATTTGGAAAGTAATTGGTTT  ${ t TGTCCTGAAGGAAACCACTGGCTTGATATTTCTGTGACTCGTGTTTGCCTTTGAAACTAGTCC}$ CCTACCACCTCGGTAATGAGCTCCATTACAGAAAGTGGAACATAAGAGAATGAAGGGGCAGA TAGCTGAGAAATTGTTGACATAAAATAAAGAATTGAAGAAACACATTTTACCATTTTGTGAA TTGTTCTGAACTTAAATGTCCACTAAAACAACTTAGACTTCTGTTTGCTAAATCTGTTTCTT AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

MKRLPLLVVFSTLLNCSYTQNCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAA NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTL SNSTLTEFVKTVNNFVQRDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT NSTDIALKVFFFDSYNMKHIHPHMNMDGDYINIFPKRKAAYDSNGNVAVAFLYYKSIGPLLS SSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAF WNYSPDTMNGSWSSEGCELTYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCSIIAGL LHYFFLAAFAWMCIEGIHLYLIVVGVIYNKGFLHKNFYIFGYLSPAVVVGFSAALGYRYYGT TKVCWLSTENNFIWSFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA LALLFLLGTTWIFGVLHVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEEYYRLFKNV PCCFGCLR

#### Signal peptide:

amino acids 1-19

#### Transmembrane domain:

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636, 648-664

#### N-glycosylation site.

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181, 188-192, 249-253, 381-385, 395-399

Glycosaminoglycan attachment site.

amino acids 49-53

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 360-364

Casein kinase II phosphorylation site.

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154, 155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329, 346-350, 608-612

Tyrosine kinase phosphorylation site.

amino acids 36-44, 669-677, 670-678

N-myristoylation site.

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394, 434-440, 480-486, 521-527

Aspartic acid and asparagine hydroxylation site.

amino acids 75-87

CTAAGCGAGGCCTCCTCCCCGCAGATCCGAACGGCCTGGGCGGGGTCACCCCGGCTGGGA GGTGTGAGTGGGTGTGTGCGGGGGGGGGGGGGGTTGATGCAATCCCGATAAGAAATGCTCGGG TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG CCGCCGCGCCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC GGCACTCACAGCCCCGCAGCGCATCCCGGTCGCCGCCCCATCGCCGG  ${\tt AGCTGCGCCGAGAGCCCCAGGGAGGTGCCC}$ ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGCCCCCTCGCCTTCTCGGACGCGGGCC CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC ACGGGCTCTCCAGCTGCTTCCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGC CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG CGTGCACAGCGTGCGGTACCTCTGCATGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA CAGAGGCTTTCTTCCACTCTCATTTCCTGCCCATGCTGCCCATGGTCCCAGAGGAGCCTG AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCCTGGAGACCGACAGCATG  ${\tt GACCCATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAG\underline{{\tt TAA}}CT}$ GAGACCATGCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTCTGCTCCCTCGA GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTCAGGAACAGGTGATCCACTCTGTA AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCC  ${ t CTGAGGCCAGTTCTGTCATGGATGCTGTCCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC}$ TTCCATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCT CCCAGGCCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC TTTTGTATATTAAAATGGAGTTTGTTTGT

MRSGCVVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWGDPIRLRHLYTSGPHGLSSCFLRI RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD MFSSPLETDSMDPFGLVTGLEAVRSPSFEK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 48-59

CCCAGAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA CTGCTGCGCTACCTGGTGGTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCCAA AGAAGACTGTTTCCTCCAGATTAGAGTGGAAGAAACTGGGTCGGAGTGTCTCCTTTGTCTAC TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG  ${\tt AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT}$ CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAACTGTGGTAGAGCTACGATGTCAAGA CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA ATCCCAGACTTGGCTCCCAAAGCACCCAACAGCTCATACACAATGAATACAAAAACTGGAACT CTGCAATTTAATACTGTTTCCAAACTGGACACTGGAGAATATTCCTGTGAAGCCCGCAATTC TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT GCTCAGAGGAAAGGCTACTTTTCAAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG  ${\tt CCGCGGCGGGGGTCACGAGGTCAGGAGTTC}{{\tt TAG}}{\tt ACCAGTCTGGCCAATATGGTGAAACCC}$ CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC TTGGGAGACAGGAGATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC 

# 

# FIGURE 24

MARRSRHRLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE WKKLGRSVSFVYYQQTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQNLEED TVTLEVLVAPAVPSCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST NSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNSVGYRRCPGKRMQVDDLNISGIIAAVVVVA LVISVCGLGVCYAQRKGYFSKETSFQKSNSSSKATTMSENVQWLTPVIPALWKAAAGGSRGQEF

#### Signal peptide:

amino acids 1-20

#### Transmembrane domain:

amino acids 130-144, 238-258

#### N-glycosylation site.

amino acids 98-102, 187-191, 236-240, 277-281

#### Casein kinase II phosphorylation site.

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

#### N-myristoylation site.

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

#### Amidation site.

amino acids 226-230

GACATCGGAGGTGGGCTAGCACTGAAACTGCTTTTCAAGACGAGGAAGAGGGAGAAAGAG AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT CATGCTGCTATTCCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA TTATATCATTAAGGAAATAGTAACCTTCTCTCTCCAATATGCATGACATTTTTGGACAATG CAATTGTGGCACTGGCACTTATTTCAGTGAAGAAAACTTTGTGGTTCTATGGCATTCATCA  ${ t TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAAGCTAAG{ t ATG}{ t AAGGACATGCCACT}$ CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT AGCTAACACACAGATTCTTCTCCTACAGACTAACAATATTGCAAAAATTGAATACTCCACAG ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT ACTGCCTGAAAAATGTCTGTCCGAACTGAGCAACTTACAAGAACTCTATATTAATCACAACT TGCTTTCTACAATTTCACCTGGAGCCTTTATTGGCCTACATAATCTTCTTCGACTTCATCTC AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT TCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACTTTAAGCCTCTTATCA ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT GGACTGGAAAACTTAGAAAGCATCTCTTTTTACGATAACAGGCTTATTAAAGTACCCCATGT GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC TACTAACAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG CCAAACCTCAAGGAAATCAGCATACACAGTAACCCCATCAGGTGTGACTGTGTCATCCGTTG GATGAACATGAACAAAACCAACATTCGATTCATGGAGCCAGATTCACTGTTTTGCGTGGACC CACCTGAATTCCAAGGTCAGAATGTTCGGCAAGTGCATTTCAGGGACATGATGGAAATTTGT CTCCCTCTTATAGCTCCTGAGAGCTTTCCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT TTCCTTTCACTGTAGAGCTACTGCAGAACCACAGCCTGAAATCTACTGGATAACACCTTCTG GTCAAAAACTCTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA GATATAAATGGCGTAACTCCCAAAGAAGGGGGTTTATATACTTGTATAGCAACTAACCTAGT GCTCTTTGAATATTAAAATAAGAGATATTCAGGCCAATTCAGTTTTGGTGTCCTGGAAAGCA AGTTCTAAAATTCTCAAATCTAGTGTTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC CATCAACTGAGTATAAAATTTGTATTGATATTCCCACCATCTATCAGAAAAACAGAAAAAAA TGTGTAAATGTCACCACCAAAGGTTTGCACCCTGATCAAAAAGAGTATGAAAAGAATAATAC CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGGATTATTGGTGTGATATGTCTTATCA AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA  ${\tt AGAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCC{f T}}$ **AA**AAACCACCAAGGAAACCTACTCCAAAAATGAAC

MKDMPLRIHVLLGLAITTLVQAVDKKVDCPRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT FPARLPANTQILLLQTNNIAKIEYSTDFPVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE ENKLTELPEKCLSELSNLQELYINHNLLSTISPGAFIGLHNLLRLHLNSNRLQMINSKWFDA LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL IKVPHVALQKVVNLKFLDLNKNPINRIRRGDFSNMLHLKELGINNMPELISIDSLAVDNLPD LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSNPIRC DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGQNVRQVHFRDMMEICLPLIAPESFPSNLNV EAGSYVSFHCRATAEPQPEIYWITPSGQKLLPNTLTDKFYVHSEGTLDINGVTPKEGGLYTC IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFV KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE YEKNNTTTLMACLGGLLGIIGVICLISCLSPEMNCDGGHSYVRNYLQKPTFALGELYPPLIN LWEAGKEKSTSLKVKATVIGLPTNMS

#### Signal sequence:

amino acids 1-22

#### Transmembrane domain:

amino acids 633-650

#### N-glycosylation site.

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583, 608-612, 624-628, 625-629

#### Casein kinase II phosphorylation site.

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

#### Tyrosine kinase phosphorylation site.

amino acids 570-579

#### N-myristoylation site.

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443, 491-497, 492-498, 634-640, 702-708

#### Cell attachment sequence.

amino acids 277-280

GCCCGGGACTGCCCAAGCCAAGCAAGAAATAATGAAGAGACACATGTGTTAG CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTAAGG ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCCTCAATGCTGC CAACGACGCTGACCTTTGTAACCTCCCTAAAAAAACTACCGATTATGCCATGCTGGTCACCA TGTTTGGCTGGTTCACTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG ACCTGATGATATTAGCACTGTGGTA<u>TAG</u>TGTCCAAACTGACTGTCATTGAGAAAGAAAAA TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCCGATGGTATATTTCTGAGT 

MNLVDLWLTRSLSMCLLLQSFVLMILCFHSASMCPKGCLCSSSGGLNVTCSNANLKEIPRDL PPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDLSDNR IQSVHKNAFNNLKARARIANNPWHCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFL NAANDADLCNLPKKTTDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKK ADEPDDISTVV

#### Signal sequence:

amino acids 1-33

#### Transmembrane domain:

amino acids 205-220

N-glycosylation site.

amino acids 47-51, 94-98

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 199-203

Casein kinase II phosphorylation site.

amino acids 162-166, 175-179

N-myristoylation site.

amino acids 37-43, 45-51, 110-116

 ${ t ACCGAGCCGAGCGGAAGGCGCGCGCGAG}$ GGCGTGAGGAGCATGCCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT ACCGCGCTGTGCTGTCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG ACGCGCCTGCTGGACCTAGGCAAGAACCGCATCAAAACGCTCAACCAGGACGAGTTCGCCAG CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCGGCG CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATC CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG ACAATGACCTCGTCTACATCTCTCACCGCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCT CATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAC TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCCTCAACCTCTCCTACAACCCCATCA GCACCATTGAGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC GGGCAGCTGGCGTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCGTGCTCAA TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTCGGCTCCTGTGGGTGTTCCGG CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCCACGTGCGCCACGCCCGAGTTTGTCCA GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACCTGCCGCCGC CCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGCCACACGGTGCAGTTT GTGTGCCGGGCCGATGGCGACCCGCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT ACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCCAACGCGGCCGAACGAC TCCATGCCCGCCCACCTGCATGTGCGCAGCTACTCGCCCGACTGGCCCCATCAGCCCAACAA GACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGGAGAGGCCCAACAGCACCCGCGCCACTG TGCCTTTCCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTTTTC  $\tt CTGGGCGTCGTCTTCTGCCTGGTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC$ AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG  ${\tt ACGCGCCCGCAAGTTCAACATGAAGATGATA}{\tt TGA}{\tt GGCCGGGGGGGGGGGGGACCCCCG}$ GGCGGCCGGGCAGGGGAAGGGGCCTGGTCGCCACCTGCTCACTCTCCAGTCCTTCCCACCTC CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCA  ${ t ATAATTCAATAAAAAGTTACGAACTTTCTCTGTAACTTGGGTTTCAATAATTATGGATTTT$ 

MQVSKRMLAGGVRSMPSPLLACWQPILLLVLGSVLSGSATGCPPRCECSAQDRAVLCHRKCF VAVPEGIPTETRLLDLGKNRIKTLNQDEFASFPHLEELELNENIVSAVEPGAFNNLFNLRTL GLRSNRLKLIPLGVFTGLSNLTKQDISENKIVILLDYMFQDLYNLKSLEVGDNDLVYISHRA FSGLNSLEQLTLEKCNLTSIPTEALSHLHGLIVLRLRHLNINAIRDYSFKRLYRLKVLEISH WPYLDTMTPNCLYGLNLTSLSITHCNLTAVPYLAVRHLVYLRFLNLSYNPISTIEGSMLHEL LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSVGNLETLILDSNPLA CDCRLLWVFRRWRLNFNRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV FVDEGHTVQFVCRADGDPPPAILWLSPRKHLVSAKSNGRLTVFPDGTLEVRYAQVQDNGTYL CIAANAGGNDSMPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIEIEYVPRKSDAGISSADAPRKFNMKMI

#### Signal sequence:

amino acids 1-41

#### Transmembrane domain:

amino acids 556-578

#### N-glycosylation site.

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345, 492-496, 505-509, 526-530, 542-546

#### Casein kinase II phosphorylation site.

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353, 607-611

#### Tyrosine kinase phosphorylation site.

amino acids 590-598

#### N-myristoylation site.

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143, 262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

# Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 32-43

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGGCGCCCTTTCGGTCAAC ATCGTAGTCCACCCCTCCCCATCCCCAGCCCCCGGGGATTCAGGCTCGCCAGCCCCAGCC  ${\tt AGGGAGCCGGCGGGAAGCGCG} \underline{{\tt ATG}} {\tt GGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCT}$ ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCA CGAGGACTCATCCCTGCAATGGTCTAACCCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA GAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCCACAGAAGCCCATCATCACTGGTT ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAAACTGTCAGTCTTCTGGGAGCAAG CCTGCAGCCCGGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCCAACCCGCAT ACAGGAAGATCCCAATGGTAAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCC TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTCGCGGCAATCCAGTCC CCCAGCAGTACCTATGGGAGAGGGGGGGGGTGTGCCACCCCTGAAGATGACCCAGGAGAGT GCCCTGATCTTCCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG CAACATGGGCAGCTACAAGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGCCCT CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCCTGCTG TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG GGCAGTCAGGAGGGACGACAAGAAGGAATATTTCATCTAGAGGCGCCTGCCCACTTCCTGC GCCCCCAGGGGCCCTGTGGGGACTGCTGGGGCCGTCACCAACCCGGACTTGTACAGAGCAA TTGCCCTCAGCCCTTTCCGTGGCTTCTCTGCATTTGGGTTATTATTATTTTTTGTAACAATCC AACAAAAAACA

MGAPAASLLLLLLLFACCWAPGGANLSQDDSQPWTSDETVVAGGTVVLKCQVKDHEDSSLQW SNPAQQTLYFGEKRALRDNRIQLVTSTPHELSISISNVALADEGEYTCSIFTMPVRTAKSLV TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRIQEDPNGK TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQRIEVLYTPTAMIRPDPPHPREGQ KLLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATSNMGSYKA YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLLIMLIFLGHYLIRHKGTYLTHEAKGSDD APDADTAIINAEGGQSGGDDKKEYFI

#### Signal sequence:

amino acids 1-20

#### Transmembrane domain:

amino acids 331-352

#### N-glycosylation site.

amino acids 25-29, 290-294

#### Casein kinase II phosphorylation site.

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

#### N-myristoylation site.

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304, 306-310, 334-340, 360-364, 385-389, 386-390

### Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

GGGGGTTAGGGAGGAATCCACCCCCACCCCCAAACCCTTTTCTTCTTCTTCCTGG CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG  ${\tt TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAA} \underline{{\tt ATG}} {\tt CTGCTTTGGATTCTGTT}$ GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAAGA TCTGTTCCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAAGGGCTTCACA CCTCACTCGACTTTTCCCTAATGAGTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTTCTGGGGGCTGCAGCTGGTGAAAAGG CTGCACATCAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA TCTGGAATATCTCCAGGCTGATTTTAATTTATTACGAGATATAGACCCGGGGGCCTTCCAGG ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCCTACCTGCCAAC GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC CTATGAGGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCCTT GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC CACCGAACAGGACTTGTGTCCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG CCCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACA GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTCGGGTTTAAAGATGAAC TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCCAAGCTCTCTAACGTGCAGGA GCTTTTCCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA TCCTCCCGGGCACTTCAATGCCATGCCCAAACTGAGGATCCTCATTCTCAACAACAACCTG CTGAGGTCCCTGCGTGGGCCTGTCGCTGCGCTCTCTAAACTCAGCCTGCACAA CAATTACTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCTTTCAAGCAGTGGGCA GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACTTCTT TAGAAAGGATTTCATGCTCCTCCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT CGCCCACGTTAACTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT GTTTGTCACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTTTATCCTGAGGAACCGAAAGC GACTCTTCCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCCACAGAGTGTATGACTG TGGCTCTCACTCGCTCTCAGAC<u>TAA</u>GACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG ATACATCCTTCCCCACCGCAGGCACCCCGGGGGGCTGGAGGGGGCGTGTACCCAAATCCCCGCG CCATCAGCCTGGATGGGCATAAGTAGATAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA GCCAGCTCGCTCTTTGCTGAGAGCCCCTTTTGACAGAAAGCCCAGCACGACCCTGCTGGAAG AACTGACAGTGCCCTCGCCCTCGGCCCCGGGGCCTGTGGGGTTGGATGCCGCGGTTCTATAC ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCCTGTGGATTAG CCCCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT TGTAAATAAGTAACTTTGACTTCTGAC

MLLWILLLETSLCFAAGNVTGDVCKEKICSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH LFLHGNSLTRLFPNEFANFYNAVSLHMENNGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQ TFLGLDDLEYLQADFNLLRDIDPGAFQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRG NRLKTLPYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTK IPGNWQIKIRPTAAIATGSSRNKPLANSLPCPGGCSCDHIPGSGLKMNCNNRNVSSLADLKP KLSNVQELFLRDNKIHSIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNY LDTLSREKFAGLQNLEYLNVEYNAIQLILPGTFNAMPKLRILILNNNLLRSLPVDVFAGVSL SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGNPWECSCTIVPFKQWAERLGSEVLMSDLKC ETPVNFFRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL VPGLLLVFVTSAFTVVGMLVFILRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG AHRVYDCGSHSLSD

#### Signal sequence:

amino acids 1-15

#### Transmembrane domain:

amino acids 618-638

#### N-glycosylation site.

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 122-126, 646-650

#### Casein kinase II phosphorylation site.

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577, 608-612, 657-661, 666-670, 693-697

#### N-myristoylation site.

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349, 354-360, 465-471, 493-499, 598-604, 603-609

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 337-348

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCCTGACCCCAGAATAACTCAGGGC TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCCTGTCGCGGCCTAAGGGAAACTGTTGGC CGCTGGGCCCGCGGGGGATTCTTGGCAGTTGGGGGGGTCCGTCGGGAGCGAGGGCGAGGGG AAGGGAGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC CGGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGC ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTC AGCACCGTGCGTGCGGGCCCGAGCTGCGCGCTGTGCTGCGCGCTCCTGCGGGCAGGCCCAGG GCCCGGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTCGCACTGGAGCGCAGGCGTTCCCACT GCACCCTGGAGAACGAGCCTTTGCGGGGGTTTCTCCTGGCTGTCCTCCGACCCCGGCGGTCTC GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCCAACGCTCCTGCACCGCGCGGAGATGCGC GGTACTCCAGGCCACCGGTGGGGTCGAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCCTGCGCCGCGCCCCGGG GCCGCCTCTAACTTGAGCTATCGCGCGCCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG TCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA TCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCCTGCCCC GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA GCAACCAGCCCCGTGCCGCAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC ACCACTTGTCCCTGAACAAGACAATTCAGTAACATCTATTCCTGAGATTCCTCGATGGGGAT CACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC CCATCAGGGAGCGTGATTTCCAAGTTTAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTT CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTTGGTGATCT TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCCTCTTCCCAGCCA AGGAAGGAGTCTATGGGCCCGCCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG CAGAGGGTGCCTTGCTGGCGAGTCCCCTCTTGGCTCTAGTGATGCA<u>TAG</u>GGAAACAGGGGA CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTAC TTGTGTAACTGACAATTTCTGCAGAAATCCCCCTTCCTCTAAATTCCCTTTACTCCACTGAG GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAA TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA TTAAAAAATAATTTCTACCAAAATGGAAAGGAAATGTTCTATGTTGTTCAGGCTAGGAGTAT ATTGGTTCGAAATCCCAGGGAAAAAAATAAAAATAAAAATTAAAGGATTGTTGAT

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAAEEACILRGGALS TVRAGAELRAVLALLRAGPGPGGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLCPAPRPGA ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCIADEIGARWDKLSGDVLCPCPG RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTSGEGQPTLGGTGVPTRRPPATA TSPVPQRTWPIRVDEKLGETPLVPEQDNSVTSIPEIPRWGSQSTMSTLQMSLQAESKATITP SGSVISKFNSTTSSATPQAFDSSSAVVFIFVSTAVVVLVILTMTVLGLVKLCFHESPSSQPR KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

#### Signal sequence:

amino acids 1-16

#### Transmembrane domain:

amino acids 399-418

#### N-glycosylation site.

amino acids 189-193, 381-385

#### Glycosaminoglycan attachment site.

amino acids 289-293

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 98-102, 434-438

#### Casein kinase II phosphorylation site.

amino acids 275-279, 288-292, 342-346, 445-449

#### N-myristoylation site.

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157, 185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469, 477-483

### Aspartic acid and asparagine hydroxylation site.

amino acids 262-274

CGGACGCGTGGGATTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG GCGCCGGTGTGAGCCAGCGCTGCCAGTGTGAGCGGCGGTGTGAGCGCGGTGCGGA GGGGCGTGTGTGCCGGCGCGCGCGCGCGTGGGGTGCAAACCCCGAGCGTCTACGCTGCCATGA CAGCAGTCCCCAGAGAGACCTGTTTTCACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT TATTGGCAGTGAAGGTTTTCCTGGAGTGTACCCTCCAAATAGCAAATGTACTTGGAAAATCA CAGTTCCCGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTGGCCG CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAAC CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC CGATATGATTATGTGGCTGTTTTAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGGAAA GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTCAGGCCAAAA AAACTGCCTACAACTACAGAACAGCCTGTCACCACCACATTCCCTGTAACCACGGGTTTAAA GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAACCATCACTCGCGATGGGAGTTTG CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTCAGCAGGCGGG CAAGAACATGAGTGCCAGGCTGACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG  ${\bf T} \underline{{\bf TAA}} {\tt CAGTGAACTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTC}$ TCTCAGTAGAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAGATGG GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTT TGCCTGCTGTCAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA  ${ t ATTTTAGAATTGAGTTGTGAAGATGTCAAAAAAAGATTTTAGAAGTGCAATATTTATAGT$ GTTATTTGTTTCACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTA 

MRGANAWAPLCLLLAAATQLSRQQSPERPVFTCGGILTGESGFIGSEGFPGVYPPNSKCTWK
ITVPEGKVVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGNKMMV
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPPAPIVSERNELLI
QFLSDLSLTADGFIGHYIFRPKKLPTTTEQPVTTTFPVTTGLKPTVALCQQKCRRTGTLEGN
YCSSDFVLAGTVITTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLRR
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

### Signal sequence:

amino acids 1-23

N-glycosylation site.

amino acids 355-359

Casein kinase II phosphorylation site.

amino acids 64-68, 142-146, 274-278

Tyrosine kinase phosphorylation site.

amino acids 199-208

N-myristoylation site.

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295, 305-311, 309-315, 320-326, 330-336

Cell attachment sequence.

amino acids 149-152

CGGACGCGTGGGCGGCCGGCCCACGGCGCCCGCGGGCTGGGCCGCTTCTT  $\tt CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAG{\color{red} ATG}{\color{blue} GCCCCCGAAGG}$ GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTCCTCAACCTCCCAGGACCTATCTGG CTCCAGCCTCTCCACCTCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTCATACCTG CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACTTTGGAG GTGGAAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT GAGTGAGGAGCTGGTGGTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCC AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGC CTTCCCTGTCCTGGGGGAACAGAGGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG GACACGAGGGGGCAGCGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC TTTGGCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACTGTTTGCAATGCAAGAAGGG CTGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG GCCTGCCTAGGCTGCATGGGGGCCAGGCCAGGTCGCTGTAAGAAGTGTAGCCCTGGCTATCA GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC AGAAGACGAGTTGGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA CGCTGGCTGAAGGGCGACTTGGTGTTCACCGCCATCTTCATTGGGGCTGTGGCGGCCATG ACTGGCTACTGGTTGTCAGAGCGCAGTGACCGTGTGCTGGAGGGCCTTCATCAAGGGCAGA**TA A**TCGCGGCCACCACTGTAGGACCTCCTCCCACCCACGCTGCCCCAGAGCTTGGGCTGCCC TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG CCTTACAGAGCAGCCCAGGTACCCAGGCCCGGGCAGACAAGGCCCCTGGGGTAAAAAGTAGC CCTGAAGGTGGATACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA ATTTCAAAAGTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCTGCTTAGGATTAGGTG GTCCTCACAGGGGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT 

MAPWPPKGLVPAVLWGLSLFLNLPGPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER TIRDNFGGGNTAWEEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLELSEELVESWWFHKQ QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG YGGEACGQCGLGYFEAERNASHLVCSACFGPCARCSGPEESNCLQCKKGWALHHLKCVDIDE CGTEGANCGADQFCVNTEGSYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDELVVLQQMFFGIIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

#### Signal sequence:

amino acids 1-29

#### Transmembrane domain:

amino acids 372-395

#### N-glycosylation site.

amino acids 79-83, 205-209

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 290-294

### Casein kinase II phosphorylation site.

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

### N-myristoylation site.

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179, 177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289, 326-332, 372-378, 395-401

### Aspartic acid and asparagine hydroxylation site.

amino acids 321-333

### EGF-like domain cysteine pattern signature.

amino acids 181-193

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA  $\tt GCACC{\color{red} ATG} CAGCCCCTGTGGCTCTGGGCACTCTGGGTGTTGCCCCTGGCCAGCCCCGGG$ GCCGCCTGACCGGGGAGCAGCTCCTGGGCAGCTGCTGCGGCAGCTGCAAAGAGGT ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTCAGCCAGAGC TTCCGAGAGGTGGCCGGCAGGTTCCTGGCGTTGGAGGCCAGCACACCTGCTGGTGTTCGG CATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGGCTCTTCCAGG CTCCAGGCTGGTGTCCACGAGAGCCGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTACAGGTGTCGGTGCAGAGG GAGCATCTGGGCCCGCTGGCGTCCGGCGCCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG CTCAGGGCGACTGTGACCCTGAAGCACCAATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG ATGTACATTGACCTGCAGGGGATGAAGTGGGCCGAGAACTGGGTGCTGGAGCCCCCGGGCTT CCTGGCTTATGAGTGTGGGCACCTGCCGGCAGCCCCGGAGGCCCTGGCCTTCAAGTGGC ATCAAGGAGGGAGGCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA  ${ t GTGCAGCTGTGCCTCGGATGGTGCGCTCGTGCCAAGGAGGCTCCAGCCA}$ TAGCCATCGAGGGACTTGACTTGTGTGTTTTCTGAAGTGTTCGAGGGTACCAGGAGAGCTG GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTAGTGAGCCCCTGAATTTGCTT CCTCTGACAAGTTACCTCACCTAATTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA  ${\tt GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT}$ GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCANTGTGTCATTGTTTACTTGTCCTGTCAC TGGATCTGGGCTAAAGTCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT AAAA

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV ALLQRSHGDRSRGKRFSQSFREVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP VPKAALHRHGRLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKCSCASDGALVPRRLQP

### Signal sequence:

amino acids 1-18

N-glycosylation site.

amino acids 158-162

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 76-80

Casein kinase II phosphorylation site.

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

N-myristoylation site.

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

Amidation site.

amino acids 74-78

TGF-beta family signature.

amino acids 282-298

 $\tt GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCG{\color{red} \underline{ATG}} GGGACAA{\color{blue} A}$ GGCGCAAGTCGAGAGAAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG CATTGGGCAGTGTTACAGTGCACTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCCGTGTGGAAGTTTGACCA AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT GGGATAGTGATGCCTACGAATCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT GTGGAGCGGAATGTGGGGGTCATCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG  ${\tt ACCTCGTCATTCCTGGTG} \underline{{\tt TGA}} {\tt GCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT}$ CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC TACACCCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC ATCCTCCTTCATGCCCTCCCTCTCCTTCCTACCACTGCTGAGTGGCCTGGAACTTGTTTAAA GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC TTCTAAGTAGACAGCAAAAATGGCGGGGGTCGCAGGAATCTGCACTCAACTGCCCACCTGGC  ${\tt TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGGTTCTGGGCTCTTTCCTTGTGTACTGAC}$ GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACTGGAGGCTGGGCGCAGTGGCTCACGCCTG TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA ACATGGAGAAACCCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC 

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW KFDQGDTTRLVCYNNKITASYEDRVTFLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPSEYTWFKDGIVMPTNPKSTRAFSNS SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVAAVLVTLI LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

### Signal sequence:

amino acids 1-27

#### Transmembrane domain:

amino acids 238-255

### N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 270-274

### Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158, 193-197, 203-207, 287-291

### N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

 ${\tt CAGCGCGTGGCGCGCTGTGGGGACAGC} \underline{{\tt ATG}} \underline{{\tt AGCGGCGGTTGGATGGCGCAGGTTGGA}}$ GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTCGGCCTCGGACTAGGCCT GGAGGCCGCGGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG GCTCGTGCCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCTCACCTGG CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGTGCAGGATTGAGCC ATGTACCCAGAAAGGGCAATGCCCACCGCCCCTGGCCTCCCCTGCCCCTGCACCGGCGTCA GTGACTGCTCTGGGGGAACTGACAAGAAACTGCGCAACTGCAGCCGCCTGGCCTAGCA GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCACTCACGTGGCGCTGCGACGGCCA CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCACCTCTCTCAGGAATGCC ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCCTCTGTCGGGAATGCCACATCCTC CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC TCAGTGCAAGCCTGGTCACCGCCACCCTCCTTTTTGTCCTGGCTCCGAGCCCAGGAGCGC CTCCGCCCACTGGGGGTTACTGGTGGCCATGAAGGGGTCCCTGCTGCTGTCAGAACAGAAGAC  $\tt CTCGCTGCCC\underline{TGA} \tt GGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA$ GGAGGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAGTGGCCCTGGAGATTGAGGGTCCC TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACCTGCCACAGCCAGAACTGAG GGGCTGGCCCCAGGCAGCTCCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

MSGGWMAQVGAWRTGALGLALLLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR TSGLCVPLTWRCDRDLDCSDGSDEEECRIEPCTQKGQCPPPPGLPCPCTGVSDCSGGTDKKL RNCSRLACLAGELRCTLSDDCIPLTWRCDGHPDCPDSSDELGCGTNEILPEGDATTMGPPVT LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAAVLSASLVTATLL LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

#### Signal sequence:

amino acids 1-30

#### Transmembrane domain:

amino acids 230-246

#### N-glycosylation site.

amino acids 126-130, 195-199, 213-217

### Casein kinase II phosphorylation site.

amino acids 84-88, 140-144, 161-165, 218-222

### N-myristoylation site.

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218, 224-230, 230-236, 263-269

### Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 44-55

### Leucine zipper pattern.

amino acids 17-39

CCCACGCGTCCGGTCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG GTTAGACTGGCGGGGGGAGGAGGCGGAGGGAAGGAAGCTGCATGAGACCCACAGA CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCCAGCG GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA TAATTCCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG TACCCCGACCTACACAATATGGTTTCATTATGTCGCGATGATGGAACGTGGAATAATCTGCC CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCCGGTG CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCGTCT GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAACTGTGGTGGAGTTTTACTGCGATCCT GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTCCTTC TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA CCACGTGGAAGATTGTGGCGTTCACGGCAACCAGTGTGCTGCTGGTGCTGCTCGTCATC CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCCAGGGGGGCCTCCCCGGAG TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG ACGAAGCTGTGAGTGCCGTTGAGTGCCTTAGGCCCCGGGTACATGGCCTCTGTGGGCCAG GGCTGCCCCTTACCCGTGGACGACCAGAGCCCCCCAGCATACCCCGGCTCAGGGGACACGGA CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA  $\tt GTTGTTCCTAAGAAAC{\color{red}{\textbf{TGA}}}TTGATTAAAAAATTTCCCAAAGTGTCCTGAAGTGTCTCTTCAA$ ATACATGTTGATCTGTGGAGTTGATTCCTTTCCTTCTTGGTTTTAGACAAATGTAAACAA AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC CTGTTTCTTCACACAGACTGATTAAAAATTAAAAGNAAAAA

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS
SNGYVNISELQTSFPVGTVISYRCFPGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLP
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT
WPSTHETLLTTWKIVAFTATSVLLVLLLVILARMFQTKFKAHFPPRGPPRSSSSDPDFVVVD
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGDTDTGPGESETCDS
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEEVASTSPGIHHAHWVLFLRN

### Signal sequence:

amino acids 1-41

#### Transmembrane domain:

amino acids 325-344

#### N-glycosylation site.

amino acids 104-108, 134-138, 192-196

### Casein kinase II phosphorylation site.

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366, 364-368, 380-384, 467-471, 468-472

#### N-myristoylation site.

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424, 478-484

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 394-405

CCCACGCGTCCGCCCCCCCCCCCCCCCCCGTGCGGTCCGTCGGTGGCCTAGAGA TGCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT  ${\tt AGCGCCCGAGTGTCGGGGCGCACCCGAGTCGGGCC}$ TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAA AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATAT CACAATTTAGGAACTGGTATGTGGATGAGCCGTCCTGCGGCGAGGTCTGCGTGGTCATG TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCCTACATGTTCCAGTGGAATGATGA CCGGTGCAACATGAAGAACAATTTCATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTT CTAGAGAAGCTGAAGGTGAGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG GAAGAAGATGCCAAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT CCTAATCCCCAGCATTCCCCTTCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAGAAGAAGCAACACACCATC TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCATTCCGAGTGTGTT CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG GTTATTAGGACATATAAAAAACTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA GAGTCTCAATAAATGTCACTTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG GCAGTGATAAAGATGGGCTGTGGAGCTTGGAAAACCACCTCTGTTTTCCTTGCTCTATACAG CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAG GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTCAAAGAATAAAATCAAATAAAGA GCAGGAAAAAAAAA

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRGGTQRPCYKVIYFHDTSRRL
NFEEAKEACRRDGGQLVSIESEDEQKLIEKFIENLLPSDGDFWIGLRREEKQSNSTACQDL
YAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLLV
VTTVVCWVWICRKRKREQPDPSTKKQHTIWPSPHQGNSPDLEVYNVIRKQSEADLAETRPDL
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES
GWVENEIYGY

#### Signal sequence:

amino acids 1-21

### Transmembrane domain:

amino acids 235-254

#### N-glycosylation site.

amino acids 117-121, 312-316

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 296-300

### Casein kinase II phosphorylation site.

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226, 299-303, 306-310, 323-327

#### N-myristoylation site.

amino acids 18-24, 37-43, 76-82, 146-152

ATCCGCAGGTTCCCGCGGACTTGGGGGCCCCGCTGAGCCCCGGGGCCCCGCAGAAGACTTGT  $\tt GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCCAGCGAGGGCCTACCACC{\color{red} ATG} ATCACTGGTGT$ AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTCGACCGCAGC CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAACTC AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT GCAGCAAATGTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTCGTTCCACTAACATTTTTCGGAATCTG CCACACTGATGAAGCAGATTCAGAAGTCTTGTATCCCAACTACCAAAGCTGCTGGAGCCTGA GGCAGAGACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT GGACAACGTGGCTGCCGAGCAGGCACACCACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC CATGGACTCTGCCACTGCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTGACCACAAATGGCCACCGTTTGCT GTTGACCTGACCATGGAACTTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCGCTGGACATGT TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA  ${ t ACTCAGGTGATGGAAGTTGAAGAGGT}{ t AAA}{ t CTGATTTATAAAAGCAGGATGTGTTGATT}$ TTAAAATAAAGTGCCTTTATACAATG

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMVQVVFRHGAR SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE GPIIIHTDEADSEVLYPNYQSCWSLRQRTRGRRQTASLQPGISEDLKKVKDRMGIDSSDKVD FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFLHILES NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW FVQLYYHGKEQVPRGCPDGLCPLDMFLNAMSVYTLSPEKYHALCSQTQVMEVGNEE

Signal sequence:

amino acids 1-23

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 218-222

Casein kinase II phosphorylation site. amino acids 87-91, 104-108, 320-324

Tyrosine kinase phosphorylation site. amino acids 280-288

N-myristoylation site.

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

Amidation site.

amino acids 216-220

Leucine zipper pattern.

amino acids 10-32

Histidine acid phosphatases phosphohistidine signature. amino acids 50-65

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT  ${\tt TAAATTTCAGCTCATCACCTTGCCTTGGTC} \underline{{\tt ATG}} {\tt GCTCTGCTATTCTCCTTGATCCTT}$ GCCATTTGCACCAGACCTGGATTCCTAGCGTCTCCATCTGGAGTGCGGCTGGTGGGGGGCCT CCACCGCTGTGAAGGCCGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGGCTGTGGAGCTGCCAGC GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG ATTGTTCACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA GTCCCAGAGGGTGTCAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGCCGCAAAGGTGG TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCTTCA GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCG AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACTGGGGAGAAAAGGAGGA AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTCGTTGCTCAGGGGAGGAG  ${\tt TGTGGCTGTCATCTGCTCAGTG}$ GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCCTGATTCTCAGGCCTTCAGAGTTGG ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT CACCATTTGTCCTGTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG TATGTGTTCAAA

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC RELGCGAASGTPSGILYEPPAEKEQKVLIQSVSCTGTEDTLAQCEQEEVYDCSHDEDAGASC ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCQTGWSLRAAKVVCRQLGCGRAVL TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG GDNLCSGRLEVLHKGVWGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

#### Signal sequence:

amino acids 1-15

### Casein kinase II phosphorylation site.

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238, 267-271, 294-298, 316-320, 336-340

#### N-myristoylation site.

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143, 180-186, 263-269, 286-292

#### Amidation site.

amino acids 196-200

### Speract receptor repeated domain signature.

amino acids 29-67, 249-287

ACTGCACTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGAC CCACGCGTCCGCGACGCGTGGGCCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG GTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCCTGCTGTTCGGCTGCCTGGG CGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGAAGGCCTACCTGCGGAATGCTG TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACAAGCCTTACTTGGTGACCTTCGACC TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC GACATACTTGTCAACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA TGTGGACAAGAGGGTCATGGAGACAAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTAC

MDFITSTAILPLLFGCLGVFGLFRLLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL VNNAGISYRGTIMDTTVDVDKRVMETNYFGPVALTKALLPSMIKRRQGHIVAISSIQGKMSI PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT AQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

#### Signal sequence:

amino acids 1-21

#### Transmembrane domain:

amino acids 104-120, 278-292

N-glycosylation site.

amino acids 228-232

Glycosaminoglycan attachment site.

amino acids 47-51

Casein kinase II phosphorylation site.

amino acids 135-139, 139-143, 253-257

Tyrosine kinase phosphorylation site.

amino acids 145-153, 146-153

N-myristoylation site.

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

Amidation site.

amino acids 265-269

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 6-17

CCCACGCGTCCGCTGGTGTTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA AAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGG<u>ATG</u>AAATTTCTTCTGGACATCCTC CTGCTTCTCCCGTTACTGATCGTCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCCTAA GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTCATACCTTTGT GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA GGCATTTCTTCCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG TTTCATAAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG TCTGTGTCCTAATTTCGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA CCTGGCAGTTTTAAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA AAGCGCAATAAGCACCTAGTTTTCTGAAAACTGATTTACCAGGTTTAGGTTGATGTCATCTA ATAGTGCCAGAATTTTAATGTTTGAACTTCTGTTTTTTTCTAATTATCCCCATTTCTTCAATA TCATTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA CCAAAATGACTTTATTAAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTGATTGCACTTAAATTTTTGT  ${ t TGAAGGACTATATCTAGTGGTATTTCACAATGAATATCATGAACTCTCAATGGGTAGGTTTC}$ ATCCTACCCATTGCCACTCTGTTTCCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA 

MKFLLDILLLPLLIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSK LVLWDINKHGLEETAAKCKGLGAKVHTFVVDCSNREDIYSSAKKVKAEIGDVSILVNNAGVV YTSDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYC SSKFAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRLMH GILTEQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMKAQ

### Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 30-34, 283-287

Casein kinase II phosphorylation site. amino acids 52-56, 95-99, 198-202, 267-271

N-myristoylation site.

amino acids 43-49, 72-78, 122-128, 210-216

 $\tt CCCACGCGTCCGCGGACGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCGCGGGCTC$  ${\tt AGGGAGGAGCACCGACCGCACCCTGAGAG} {\tt ATG} {\tt GTTGGTGCCATGTGGAAGGTGATTG}$ TTTCGCTGGTCCTGTTGATGCCTGGCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCCAGGACTGAACATGA AGAGTTATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTC TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG AGGTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCACAAGTAACATGA CCTTGCGTGACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT AGCACGGGATTTATACAGTGCACTAATTCAGTTTTTCCAGATATTTCCTGAATATAAAAATA ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT TGTTGGATGAGAAGCAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC AGGAAGCAGAACTGGTTTGAGGCCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGT GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCACTCCCAGAGGTGAGACAA GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACTATAGTTGAAAAGTACTTGCGAGA AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGACTTCCATCAGGTAATTATTC GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA  ${\tt TTCATTTATGGAAAAGGATGGGATCCTTATGTTGGA} {\tt TAAA} {\tt ACTACCTTCCCAAAAGAGAACAT}$ CAGAGGTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAA TTATCTTTTCATATCTGCAAGATTTTTTTCATCAATAAAAATTATCCTTGAAACAAGTGAGC TTTTGTTTTTGGGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA AATTTTAGGGTCTTGAATAGGAAGTTTTAATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTG TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT TAGTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA AAAATATTATATATAAAAGTAAAAAAAAAA

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDSGQPLFLTPYIEAGKIQKGRELSL VGPFPGLNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH GPYVVTSNMTLRDRDFPWTTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQF FQIFPEYKNNDFYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVTG CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT EIMNNYKVLIYNGQLDIIVAAALTERSLMGMDWKGSQEYKKAEKKVWKIFKSDSEVAGYIRQ AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

### Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 81-85, 132-136, 307-311, 346-350

Casein kinase II phosphorylation site.

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352, 353-357, 424-428

Tyrosine kinase phosphorylation site.

amino acids 423-432

N-myristoylation site.

amino acids 22-28, 110-116, 156-162, 232-238

Serine carboxypeptidases, serine active site.

amino acids 200-208

Crystallins beta and gamma 'Greek key' motif signature.

amino acids 375-391

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTTGGCTACAACAT  ${\tt TTTTCCCTTACAAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT}$ CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG  ${ t TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC}$ CTGGGGGAGGGCCTGCCTAACAAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT  ${\tt TGAACTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGCCAGAAA}$ TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA GACCAAAATAGATAACAGGATTCCCTGAACATTCCTAAGAGGGGAGAAAGTATGTTAAAAATA GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCTGGGTC AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTTGGTCTGACCACTCTGCCTTGTGTTTT CCTCACCGCCCCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT CCACCAAGGGACGGGGCCGTCTATGTGGGGGCCCATCAACCGGGTCTATAAGCTGACAGGCA ACCTGACCATCCAGGTGGCTCATAAGACAGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG CCCCTCATCGTGCAGCCTGCAGCGAAGTGCTCACCCTCACCAACAATGTCAACAAGCTGCT CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGGTCTGCA AGCTGCTGCGGCTGGATGACCTCTTCATCCTGGTGGAGCCATCCCACAAGAAGGAGCACTAC CTGTCCAGTGTCAACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA  ${\tt TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA}$ GCCGGAAGCTGCCCCGAGACCCTGAGTCCTCAGCCATGCTCGACTATGAGCTACACAGCGAT TTTGTCTCCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT CTTCTACATCTACGGCTTTGCTAGTGGGGGCCTTTGTCTACTTTCTCACTGTCCAGCCCGAGA CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG CTCTGCAAGGATGACCCCAAGTTCCACTCATACGTGTCCCTGCCCTTCGGCTGCACCCGGGC CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC AGGCCTTCAATATCACCAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG CAGTATCACCACCCGCCCGATGACTCTGCCCTGTGTGCCTTCCCTATCCGGGCCATCAACTT GCAGATCAAGGAGCGCCTGCAGTCCTGCTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC  ${ t TGCTGGGGAAGGACGTCCAGTGCACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA}$ CTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC ATTCACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAG  ${\tt GCAACTTTATTTTCTTGGGGAACAAAGG}{\tt TGA}{\tt AATGGGGAGGTAAGAAGGGGTTAATTTTGTG}$ ACTTAGCTTCTAGCTACTTCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA TTTCAATATTTCCCAAACTTTAAGAAAAACTTTAAGAAGGTACATCTGCAAAAGCAAA

MGTLGQASLFAPPGNYFWSDHSALCFAESCEGQPGKVEQMSTHRSRLLTAAPLSMEQRQPWP
RALEVDSRSVVLLSVVWVLLAPPAAGMPQFSTFHSENRDWTFNHLTVHQGTGAVYVGAINRV
YKLTGNLTIQVAHKTGPEEDNKSRYPPLIVQPCSEVLTLTNNVNKLLIIDYSENRLLACGSL
YQGVCKLLRLDDLFILVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY
FPTLSSRKLPRDPESSAMLDYELHSDFVSSLIKIPSDTLALVSHFDIFYIYGFASGGFVYFL
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP
GDSLAQAFNITSQDDVLFAIFSKGQKQYHHPPDDSALCAFPIRAINLQIKERLQSCYQGEGN
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG
YSVVFVGTKSGKLKKVRVYEFRCSNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

#### Signal sequence:

amino acids 1-32

#### Transmembrane domain:

amino acids 71-87

### N-glycosylation site.

amino acids 130-134, 145-149, 217-221, 381-385

### Casein kinase II phosphorylation site.

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387, 384-388, 471-475, 481-485, 530-534

#### N-myristoylation site.

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

 ${\tt CGTGGAGCCTCCGGCGGAGGCCGGCACGCTGGGACTCCTGCTGGTCGTCTTGGGCTTCCTGGTGCTCCC}$  ${\tt GCAGGCTGGAGCACCCTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGGCTGGA}$ AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCTTCGTCCTGATGGCCGCAGAGATCG GGCTGTGGGTGATTCTGCGTCCAGGCCCCTACATCTGCAGTGAGATGGACCTCGGGGGCTTGCCCAGCTGGCTAC GTTCCTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACTGC TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGATGGAGT  ${\tt ACTGGACGGGGTGGTTTGACTCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT}$ CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG GAGCCATGCACTTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG  ${\tt ATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCCCCACCTG}$ ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTACCTGTCTCTGTGGGACGCCCTCAAGTACC  ${\tt TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCT}$ TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGG TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT ACACCGTGCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGGAGAATATTGATGACCAGCGCAAAG GCTTAATTGGAAATCTCTATCTGAATGATTCACCCCTGAAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA GCTTCTTTCAGAGGTTCGGCCTGGACAAATGGNGTTCCCTCCCAGAAACACCCCACATTACCTGCTTTCTTCGG  ${\tt GTAGCTTGTCCATCAGCTCCACGCCTTGTGACACCTTTCTGAAGCTGGAGGGGCTGGAGAAGGGGGTTGTATTCA}$  ${ t TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCCAGAAGACGCTTTACCTCCCAGGTCCCTGGTTGA$  ${\tt ACCTGGGCAGGAACCAGTACATTAAG\underline{\textbf{TGA}}} {\tt GCGGTGGCACCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTC}$  $\tt CTCTTGACCTGAAGCCTGGTGGCTGCCCCACCCCTCACTGCAAAAGCATCTCCTTAAGTAGCAACCTCAGGG$ ACTGGGGGCTACAGTCTGCCCCTGTCTCAGCTCAAAACCCTAAGCCTGCAGGGAAAGGTGGGATGGCTCTGGGCC  ${\tt TGGCTTTGTTGATGATGGCTTTCCTACAGCCCTGCTCTTGTGCCGAGGCTGTCGGGGTGTGTCTCTAGGGTGGGAGC}$  $\tt TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT$  $\tt TTTATCCCCGAAATCCTGGGTGTCACCAGTGTAGAGGGTGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTT$ TCCCTTCCCACTCGCTGCTTCCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCCTGCGTCTTCC  ${\tt CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCCTGGCAGAAGCCATGTCTGCA}$ AGGGAGGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA GGAGGACAGAAGGCCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG  ${\tt ACAGAAGGCCCAGCTCAGTGGCCCCCGCTCCCCACCCCCACGCCCGAACAGCAGGGGCAGAGCAGCCCTCCTTC}$ GAAGTGTGTCCAAGTCCGCATTTGAGCCTTGTTCTGGGGCCCAGCCCAACACCTGGCTTGGGCTCACTGTCCTGA GTTGCAGTAAAGCTATAACCTTGAATCACAA

MTTWSLRRPARTLGLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW
IFGGSIHYFRVPREYWRDRLLKMKACGLNTLTTYVPWNLHEPERGKFDFSGNLDLEAFVLMA
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ
YKRGGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLLTSDNKDGLSKGIVQGVLAT
INLQSTHELQLLTTFLFNVQGTQPKMVMEYWTGWFDSWGGPHNILDSSEVLKTVSAIVDAGS
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAKYMKLRDFFGSISGIP
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYE
TSITSSGILSGHVHDRGQVFVNTVSIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN
IDDQRKGLIGNLYLNDSPLKNFRIYSLDMKKSFFQRFGLDKWXSLPETPTLPAFFLGSLSIS
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSSGINQVIVFEETMAGPA
LQFTETPHLGRNQYIK

### Signal sequence:

amino acids 1-27

Casein kinase II phosphorylation site.

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

#### N-myristoylation site.

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315, 320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG GGAGCAAAGCCGGGCTCGGCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC  ${\tt CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCC} {\tt ATG}{\tt GCTCCC}$ AAGAAGCTGTCCTGCCTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA GGCAGACACTCGGTCGTTCGTAGTGGATAGGGGTCATGACCGGTTTCTCCTAGACGGGGCCC CGTTCCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCCGAC CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACTA CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTC AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC TACAGAGCCTGTGACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGG  ${f A}{f G}{f A}{f A}{f A}{f G}{f A}{f C}{f T}{f C}{f C}{f$ GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCCTGCTT CTGGGGCCAGAATCACTCCACACGGTCTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGC TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC TATATCTGAAGCAGGGGACCCCACACCTAAGCTTTTTGCTCTTCGAGATGTCATCAGCAAGT TCCAGGAAGTTCCTTTGGGACCTTTACCTCCCCGAGCCCCAAGATGATGCTTGGACCTGTG ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCCTAGACTTGCTTTGCCCCCGTGGGCCCAT TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC GAACCTATATGACCCATACCATTTTTGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC CATGACCGTGCCTATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG AGACAAACTATTTTTGACGGGGAAACTGGGGTCCAAACTGGATATCTTGGTGGAGAACATGG GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG GGGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG GTGGTTTCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG ACCAAGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCCGGTACTGGACAAAGCAGGGGCC ACAACAGACCCTCTACGTGCCAAGATTCCTGCTGTTTCCTAGGGGAGCCCTCAACAAATTA CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACACTGAG  ${ t TGCCTCTGAACCAATGGAGTTAAGTGGGCAC{ t TGA}}$   ${ t AAGGTAGGCCGGGCATGGTGGCTCATGC}$ CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTCAGGACTTCAAGA CCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC AGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA GACACTCCATCTCAAAAAAAAAAAA

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDRGHDRFLLDGAPFRYVSGSLHYFRVPRVL
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE
YGSYRACDFSYMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF
TLLRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMYMFHGGTNF
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPPSPKMML
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPN
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGDTFLYL
PGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

#### Signal sequence:

amino acids 1-27

N-glycosylation site.

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 4-8

Casein kinase II phosphorylation site.

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554, 603-607, 644-648

Tyrosine kinase phosphorylation site.

amino acids 191-198

#### N-myristoylation site.

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233, 231-237, 274-280, 296-300, 307-313, 447-453, 484-490

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC ACCCACAAT<u>ATG</u>GCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT GCGTTCCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG AAAAACTCAGGCAGCACATTTCACGCAACGCCCAGGACAAGCAGGAGTTGCATCTGTTCATG CTGTCGGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAACT AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC TCTGCCACTGCCAAAAGTTGAACAGACTGCTTTTAGCTTTCTTCGCGATCACTTGAGA CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACTCCAGA ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT AAAACTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT AGTTTACAGAAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC CAAAACAATTGTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC CTTGGACCGCCTGCCAGCCGGCCTGGGCCAGTGTCGGATGCTCAAGAAAAGCGGGCTTGTTG TGGAAGATCACCTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA AATATTCCCTTTGCAAATGGGATT<u>TAA</u>ACTAAGATAATATATGCACAGTGATGTGCAGGAAC AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG ATACATCTTTTAAAATAAAACAGAGAGGGTGCATAGAAGGCTGATAGAAGACATAACTGAAT GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT TTCTTACTAAAAAAAAAAAAAAA

MAYMLKKLLISYISIICVYGFICLYTLFWLFRIPLKEYSFEKVREESSFSDIPDVKNDFAFL
LHMVDQYDQLYSKRFGVFLSEVSENKLREISLNHEWTFEKLRQHISRNAQDKQELHLFMLSG
VPDAVFDLTDLDVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLESLRELRHLKILHVKSNLTKVPSN
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAELELQNCELERIPHAIFSLSNLQELDLKS
NNIRTIEEIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLESLYFSNNKLESLPVAVFSLQ
KLRCLDVSYNNISMIPIEIGLLQNLQHLHITGNKVDILPKQLFKCIKLRTLNLGQNCITSLP
EKVGQLSQLTQLELKGNCLDRLPAQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP
FANGI

### Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 241-245, 248-252, 383-387

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 326-330

Casein kinase II phosphorylation site.

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

Tyrosine kinase phosphorylation site.

amino acids 349-355, 375-381

N-myristoylation site.

amino acids 78-84, 124-130, 212-218, 392-398

TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTTTACACACATAAGGAT GCTCAGTGCTTGCTTGCACTTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTCGTGCTCTGTCTTTACTTCAAAATACACAAC GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCCAGACAAGGTGTGGTGGGCCAAG AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCCTGCCCTGCAGTGCTGTGAAGGATATAGAATGTGTGCC AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAAA GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG CTCCCCATCAGTTTCATGGAAAATAACTCAGTGCCTGCTGGGAACCAGCTGCTGGAGATCCCTACAGAGAGCTTC CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCTGGAACTGACGTTTCCCTGGAGGTGTCCAGAAA GCTGATGTAACACAGAGCCTATAAAAGCTGTCGGTCCTTAAGGCTGCCCAGCGCCTTGCCAAA<u>ATG</u>GAGCTTGTA  $\overline{\mathsf{AGAAGGCTCATGCCATTGACCCTCTTAATTCTCTCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT}$ GCAAGCTGCACAGTCAGTCTAGGGGGGTGCCAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC AGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACAAAAGCATCAGAATTATCTTTTCCTATGTCCAGCTT GATCCAGATGGAAGCTGTGAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGG CAAGTCTGCAGTAAAAACGACTATGTTCCTGTATTTGAATCATCAGCAGTACATTGACGTTTCAAATAGTTACT GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTACTTCTCTCCTAACATCTCTATTCCAAACTGT GGCGGTTACCTGGATACCTTGGAAGGATCCTTCACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAAACTAAACTTCAAAGAGATTTTCCTAGAAATAGAC AAACAGTGCAAATTTGATTTCTTGCCATCTATGATGGCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT GGCCGTGTGACTCCCACCTTCGAATCGTCATCAAACTCTCTGACTGTCGTGTTGTCTACAGATTATGCCAATTCT TACCGGGGATTTTCTGCTTCCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA GACCCAACTTGCAGACCAAAATTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGGAGATAATATACATAACA GAAGATGATGTAATACAAAGTCAAAATGCACTGGGCAAATATAACACCAGCATGGCTCTTTTTGAATCCAATTCA TTTGAAAAGACTATACTTGAATCACCATATTATGTGGATTTGAACCAAACTCTTTTTGTTCAAGTTAGTCTGCAC ACCTCAGATCCAAATTTGGTGGTGTTTCTTGATACCTGTAGAGCCTCTCCCACCTCTGACTTTGCATCTCCAACC TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAAACTTGTAAGGTGTATCCCTTATTTGGACACTATGGGAGA  ${\tt TTCCAGTTTAATGCCTTTAAATTCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAAGTTTTGATATGTGAT}$ AGCAGTGACCACCAGTCTCGCTGCAATCAAGGTTGTGTCTCCAGAAGCAAACGAGACATTTCTTCATATAAATGG AAAACAGATTCCATCATAGGACCCATTCGTCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAGGATTTCAGCAT GAAACACATGCGGAAGAAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTTCCTTCATGGTTCTAGCTCTG AATGTGGTGACTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG CAGAACTAT<u>TAA</u>CTAACAGGTCCAACCCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGCTACCTCGT GGCTACACATATTATGAATAAATGAGGAAGGGCCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

MELVRRLMPLTLLILSCLAELTMAEAEGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI ERPENKSIRIIFSYVQLDPDGSCESENIKVFDGTSSNGPLLGQVCSKNDYVPVFESSSTLT FQIVTDSARIQRTVFVFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV EKDYKIKLNFKEIFLEIDKQCKFDFLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVLS TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIISKSYLEAFNSNGNNLQLKDPTCRP KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYL QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDSIIGPIRLKRDRSASGNSGFQHETHA EETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

### Signal sequence:

amino acids 1-24

#### Transmembrane domain:

amino acids 571-586

#### N-glycosylation site.

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374, 394-398, 419-423

### Casein kinase II phosphorylation site.

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383, 408-412, 463-467, 520-524, 556-560

### Tyrosine kinase phosphorylation site.

amino acids 172-180, 407-415, 407-416, 519-528

### N-myristoylation site.

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

### Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCG  ${\sf GGAC}$ TGCTGCCGCCGCCGTGCCCTGCCCACAGCGCCACGCGCTTCGACCCCACCTGGGAGTCC CTGGACGCCCGCCAGCTGCCCGCGTGGTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG GGGAGTGTTTTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAAGGAAA AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTCAAATATGAAGAT TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTCAGGC CTCTGGTGCCAAATACATTGTCTTAACTTCCAAACATCATGAAGGCTTTACCTTGTGGGGGT CAGAATATTCGTGGAACTGGAATGCCATAGATGAGGGGCCCCAAGAGGGACATTGTCAAGGAA CTTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTTGGACTGTACTATTCCCTTTTTGA ATGGTTTCATCCGCTCTTCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT ATATAATGAAAGCCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA CATAAATGGGAAAACTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAGCTGG AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG GAGGAAATCTTTTGATGAATATTGGGCCCACACTAGATGGCACCATTTCTGTAGTTTTTGAG GAGCGACTGAGGCAAGTGGGGTCCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA TACCTGGCGATCCCAGAATGACACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAG AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCCACATCAGGACAGCTGTTCCTTGGCCAT CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAACTG GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATGC CGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATC<u>TAA</u>AGTGCAGCAGAGTGGCTG ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA AAGCAATGTAAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTTCCCCACTA AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC TCTTCACATTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCACATTATAGTAG CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT TTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG **AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG** GTACAGTAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG GTTTTTCTTTCAATTATAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTAATT TTTAAAACCTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAAACATAAACTCATTGTGCAA ATGTAA

MRPQELPRLAFPLLLLLLLLPPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG
VFSVPSFGSEWFWWYWQKEKIPKYVEFMKDNYPPSFKYEDFGPLFTAKFFNANQWADIFQAS
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLFEW
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI
SLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

### Signal sequence:

amino acids 1-28

N-glycosylation site.

amino acids 171-175, 239-243, 377-381

Casein kinase II phosphorylation site.

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319, 375-375

Tyrosine kinase phosphorylation site.

amino acids 361-369, 389-397

N-myristoylation site.

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

Leucine zipper pattern.

amino acids 410-432

Alpha-L-fucosidase putative active site.

amino acids 283-295

TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACG<u>ATG</u>GCCAGGTGCTTCAGCCTGGTGTTG CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCTTGTGAGCAAAAAGGCGAACCAGC AGCTGAATTTCACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAG GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTTGGAGA TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGCAGCCTATTGTTACAACTCATCTGAT ACTTGGACTAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA AACTGCAACACAACAACAGAATTTATTGTCAGTGACAGTACCTACTCGGTGGCATCCCCTT ACTCTACAATACCTGCCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG AGAAAAAATTGATTTGTCACAGAAGTTTTTATGGAAACTAGCACCATGTCTACAGAAAC TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGGAGGTGTCC CCACGGCTCTGCTAGTGCTTCTCTCTTTTTGGTGCTGCAGCTGGTCTTGGATTTTGC CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAAACTACCGTGCGATGCCTGGAA TTACCCTGCCCCAGCTGGGGAAATCAAAAGGGCCCAAAGAACCAAAGAAGAAGTCCACCCTT GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC CCTTCTCCTTATTGTAACCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCC TTTCTAGCCTGGCTATGTCCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAA GGACCTAAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGAC CCTTTCTTCAGCTCTGAAAGAGAAACACGTATCCCACCTGACATGTCCTTCTGAGCCCGGTA AGAGCAAAAGAATGGCAGAAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACT GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA ATCACTGTTTAGAACACACACTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCT AGGAAATATACTTTTACAAGTAACAAAAATAAAAACTCTTATAAATTTCTATTTTATCTGA GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT TGAATATTATTCCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTTTCA GTTTTGATATTTCTAGCTTATCTACTTCCAAACTAATTTTTATTTTTGCTGAGACTAATCTT ATTCATTTCTCTAATATGGCAACCATTATAACCTTAATTTATTATTAACATACCTAAGAAG TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAACAAATGTATCACTA GCCCTCCTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATTTGTGACAAAAAATTAA AGCATTTAGAAAACTT

MARCFSLVLLLTSIWTTRLLVQGSLRAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACR LLGLSLAGKDQVETALKASFETCSYGWVGDGFVVISRISPNPKCGKNGVGVLIWKVPVSRQF AAYCYNSSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPP APASTSIPRRKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFF GAAAGLGFCYVKRYVKAFPFTNKNQQKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSP SKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

AGATGGCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTCGGTGCCGCGACTTTCACGATGG CTCGCCCAACCTTACTACCTTCTGTCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA ACTGCCGCCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG ACTGGAGAGAGGGGGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGAAGAACCGCAGA TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCC TGATGACGTGCAAACCCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAA ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA TTGGTCTAATGACTGCCAATCATTTGCCCCTATCTATGCTGACCTCTCCCTTAAATACAACT GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC AAAGTGAGCACATCACCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA GGAGGCAATGCGGCCGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG AGGAGAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAAACTATCA AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACCAGTGTCAGA  ${\tt TGGGGAAAACAAGAAGGATAAA} \underline{{\tt TAA}} {\tt GATCCTCACTTTGGCAGTGCTTCCTCTCTGTCAATT}$ CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG TCATAGGACGGAGGGGAAATGGTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC AGCTATTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAG TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC TTCATTAAAAGTATAAGCCTAACTTTGTCGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG TTTTTATCATTGAAGACAATATTGAACAACCCCCTATTTTGTGGGGGATTGAGAAGGGGTGAA TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATT TCCTTTTG

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRKLPPLCHGLPTQREDGNPCDFD WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

### Signal sequence:

amino acids 1-48

#### Transmembrane domain:

amino acids 111-125

N-glycosylation site.

amino acids 165-169, 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 154-158, 265-269

Casein kinase II phosphorylation site.

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

N-myristoylation site.

amino acids 188-194, 225-231

Myb DNA-binding domain repeat signature 1.

amino acids 244-253

GGACAGCTCGCGGCCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG  $\tt GGGCCCCAGCCTGGCCCCGGGTCACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA$ TTGCTCCTGCTGCCCGGCTCCTACGGACTGCCCTTCTACAACGGCTTCTACTACTCCAACAG CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG TGGTGGAGACACCCGAGGAGACCCTGTTCACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAGCT GTCGGAGAACGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCATTGACGGGCTGGA GGATGAAAGCGGTCTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCCTTACCAGTCCCCCA ACGGGCGCTACCAGTTCAACTTCCACGAGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGAGGGCCTGGACTGGTGCAACGCGGG CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCCGGCAGCCCTGCGGTGGCC GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCCTGAGAAGCT GACGCTGACAGGGCCAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC GGCAGCGTCCGCTACCCTGTGGTTCACCCGCATCCTAACTGTGGGCCCCCAGAGCCTGGGGT CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGTTTACTGCTACCGCCAGC CGTTTTCCCTTGTGGGGTTGGAGCCATTTTAACTGTTTTATACTTCTCAATTTAAATTTTCT TTAAACATTTTTTACTATTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG GATGCCCCACTCCAGGAATCATGCTTGCTCCCCTGGGCCATTTGCGGTTTTTGTGGGCTTCTG GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTTAAGGAGGTTGGTGCCCAGAGTGGGC GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCCT CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGGAGCTGGGCCTTTGGGCCTCTCCTGCC CACCTCTACTTCTCTGTGAAGCCGCTGACCCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG GGAGGGCCCTCAGGTGTGTACTTTGGACAATAAATGGTGCTATGACTGCCTTCCGCCAA 

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFTYQ GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQD KEHDVSLEIQDLRLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR HRRLHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR CDAGWLADGSVRYPVVHPHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

Signal sequence:

amino acids 1-17

Casein kinase II phosphorylation site.

amino acids 29-33, 53-57, 111-115, 278-282

Tyrosine kinase phosphorylation site.

amino acids 137-145

N-myristoylation site.

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

 ${\sf GGAGAGCGGAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC}$ TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC CAGGGGAGGTGCACCAGGCGCCCCCTGAGCGACGCTCCCCATGATGACGCCCACGGGAA CTTCCAGTACGACCATGAGGCTTTCCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCA CCCCAGAGGAAAGCCAGGCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGGAC GGCGACGCTGGGTGTCGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA CATACGGGACTCGGTGAGCGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCGGTGAAGAATTTCATGAC GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGC CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGGGCTGACAGCCTTCCTGCACCCCGAGG AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAA GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA GGAGCCGGCGTGGGTGCAGACGGAGAGGCAGCTTCCGGGACTTCCGGGATCTGAACAAGG ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCTGCCCAGGACCAGCCC CTGGTGGAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA AGCGGAAATCCTGGGTAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG ACCTGACCCGGCACCACGATGAGCTGTGAGCACCGCGCACCTGCCACAGCCTCAGAGGCCCG GCAGATGCAGTCCCAGGCATCCTCCTGCCCCTGGGCTCTCAGGGACCCCCTGGGTCGGCTTC TGTCCCTGTCACACCCCCAACCCCAGGGAGGGGCTGTCATAGTCCCAGAGGATAAGCAATAC CTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA GAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT GGCCCCAGCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA ΑΑΑΑΑΑΑΑΑΑΑΑ

MMWRPSVLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDAPHDDAHGNFQYDHEAFLGR EVAKEFDQLTPEESQARLGRIVDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT YDTDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGDSMATRE ELTAFLHPEEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEPAWVQTERQQ FRDFRDLNKDGHLDGSEVGHWVLPPAQDQPLVEANHLLHESDTDKDGRLSKAEILGNWNMFV GSQATNYGEDLTRHHDEL

Signal sequence:

amino acids 1-20

N-glycosylation site.

amino acids 140-144

Casein kinase II phosphorylation site.

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293, 291-295, 298-302

N-myristoylation site.

amino acids 263-269, 311-317

Endoplasmic reticulum targeting sequence.

amino acids 325-330

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG GCGGCGGGCGCGGGTGCGAGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTCGCTCCCAG CCTGTCTGTCGTCGTTTTGGCGCCCCCCGCCTCCCCGCGGTGCGGGGTTGCACACCGATCCTG GGCTTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG  $\tt CGGGTCGTCTGTGTCTCTCTCTCTGCGCCGCGCGCGGGGATCCGAAGGGTGCGGGGCTCT$ GAGGAGGTGACGCGCGGGGCCTCCCGCACCCTGGCCTTGCCCGCATTCTCCCTCTCTCCCAG GTGTGAGCAGCCTATCAGTCACCATGTCCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG TGTCTGCTGCTGCCGGGGCCCGCGGGCAGCGAGCGCGCTCCCATTGCTATCACATG TTTTACCAGAGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCCAGGGGGCTGCC CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC TGGTCGAGAAAACTATTCCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT GGTCTGCTTCTTTCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACTGG GCCGATTTAATTTACAGAAGAATTTTGTTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACA GAAGGACCACATGTGGGCCTTGTTCAAGCCAGTGAACATCCCAAAATAGAATTTTACTTGAA AAACTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACAT CGAGGAAGCAGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA CGGAATAATGGCTTCTTCTCTTACCACATGCCCAACTGGTTTGGCACCACAAAATACGTAAA GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTTGGAGATAGCAATTTCCGCCTC ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACCA AAGAGAATGTCCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT GATGCCATTTCCTTCACTGTTAGAAATGTGTTTTGGCCCTATAAGGGAGAGCCCCAACAAGAA CTTCCTAGTAATTGTCACAGATGGGCAGTCCTATGATGATGTCCAAGGCCCTGCAGCTGCTG CACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG  ${ t AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTCACAAGAGAGTTCACAGGATT}$  ${ t AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGTTTCTTAGAATCCCAGCAA{ t T}}$ <u>AA</u>TGGTAACATTTTGACAACTGAAAGAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT CTCATAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAAACTATTAAGTATGTCAAC AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT ACTTTGTTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA GATAATGTGGATTAAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA 

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPGGCPLEEFSVY
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVT
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKN
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLKNFTSAKDVLFAIKEVGFRGGNSNTGKAL
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG
MVQDVTFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITI
FSVGVAWAPLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

### Signal sequence:

amino acids 1-24

N-glycosylation site.

amino acids 100-104, 221-225

Casein kinase II phosphorylation site.

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424, 425-429, 478-482, 528-532

N-myristoylation site.

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211, 239-245, 240-246, 261-267, 403-409, 442-448, 443-449

Amidation site.

amino acids 145-149

GGCGGCCTCCCGGCGGAGCGAGCAGATCCAGTCCGGCCCGCAGCGCAACTCGGTCCAGTCG TGCTGGCGGCGGCGCCCCACGGCCCCCGCGCCCGACGGCGACCTCGGCTCCAGTC AAGCCCGGCCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA GGTTGAGGAACTGATGGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGG CAGAAGAAGCTGCTGAAAGCATCATCAGAAGTGAACCTGGCAAACTTACCTCCCAGCTAT TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTG TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGCCCAGC TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG TGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG GAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCACAGCCACAGCCTG GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG AGGACCTGGAGAGGACCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCT GCACTGCTGGGAGGGGAAGAGATT**TAG**ATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA TCTTCTTCCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTCAGC TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG ACAGCCGTTTGTTCTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGG AGTCTCCCTCTGATTGGTTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAA CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTCACCCTGCATTACATGTGTTTATTCATCC AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTC CCTCTCTCAGCACAGCCTGGGGAGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCCACACCAGCCTTGGTGCCACCAA AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG GTCAAACTAATTCTCACATCCCTCTAAAAGTAAACTACTGTTAGGAACAGCAGTGTTCTCAC AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT TGCATTAGTÁACTTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA GTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGCAAAATCACTTAGCAGCAAC TGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTCAGGTGTCA TGGACTGTTGCCACCATGTATTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA TAAGCATGCTTTCTTTGAGTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC 

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ
HKLRSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRGQRMLCTRDSECCG
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEEI

### Signal sequence:

amino acids 1-19

#### N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

### Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316, 327-331

### N-myristoylation site.

amino acids 202-208, 217-223

### Amidation site.

amino acids 140-144

TAAAGCCTTAAGGACAGGCCTGGCCATTACCTCTGCAGCTCCTTTGGCTTGTTGAGTCAAAAAAACATGGGAGGG CCAGGCACGGTGACTCACACCTGTAATCCCAGCATTTTGGGAGACCGAGGTGAGCAGATCACTTGAGGTCAGGAG TTCGAGACCAGCCTGGCCAACATGGAGAAACCCCCCATCTCTACTAAAAATACAAAAATTAGCCAGGAGTGGTGGC AGGTGCCTGTAATCCCAGCTACTCAGGTGGCTGAGCCAGGAGAATCGCTTGAATCCAGGAGGCGGAGGATGCAGT GGGGTAGATACTGCTTCTCTGCAACCTCCTTAACTCTGCATCCTCTTCTTCCAGGGCTGCCCCTGATGGGGCCTG GCAATGACTGAGCAGGCCCAGCCCCAGAGGACAAGGAAGAAGGCATATTGAGGAGGGCAAGAAGTGACGCCCG GTGTAGAATGACTGCCCTGGGAGGGTGGTTCCTTGGGCCCTGGCAGGGTTGCTGACCCTTACCCTGCAAAACACA  ${\tt AAGAGCAGGACTCCAGACTCTCCTTGTGAATGGTCCCCTGCCCTGCAGCTCCACC} \underline{{\tt ATG}} \underline{{\tt AGGCTTCTCGTGGCCCC}}$  $\overline{\text{ACTCTTGCTAGCTTGGGTGGCCACTGCCACTGTGCCCGTGGTACCCTGGCATGTTCCCTGCCCCCCTCA}$ GTGTGCCTGCCAGATCCGGCCCTGGTATACGCCCCGCTCGTCCTACCGCGAGGCTACCACTGTGGACTGCAATGA CCGTGTGGACCAGAGTGAGCTGGGCTACCTGGCCAATCTCACAGAGCTGGACCTGTCCCAGAACAGCTTTTCGGA TGCCCGAGACTGTGATTTCCATGCCCTGCCCCAGCTGCTGAGCCTGCACCTAGAGGAGAACCAGCTGACCCGGCT GGAGGACCACAGCTTTGCAGGGCTGGCCAGCCTACAGGAACTCTATCTCAACCACAACCAGCTCTACCGCATCGC  ${\tt CCGCTGGTTTGAAATGCTGCCCAACTTGGAGATACTCATGATTGGCGGCAACAAGGTAGATGCCATCCTGGACAT}$ GAACTTCCGGCCCTGGCCAACCTGCGTAGCCTGGTGCTAGCAGGCATGAACCTGCGGGAGATCTCCGACTATGC CCTGGAGGGGCTGCAAAGCCTGGAGAGCCTCTCCTTCTATGACAACCAGCTGGCCCGGGTGCCCAGGCGGGCACT CAACATGCTGCACCTTAAGGAGCTGGGACTGAACAACATGGAGGAGCTGGTCTCCATCGACAAGTTTGCCCTGGT GAACCTCCCGAGCTGACCAAGCTGGACATCACCAATAACCCACGGCTGTCCTTCATCCACCCCCGCGCCTTCCA CCACCTGCCCCAGATGGAGACCCTCATGCTCAACAACAACGCTCTCAGTGCCTTGCACCAGCAGACGGTGGAGTC GGTCCGTGAGGTGCCCTTCCGGGAGATGACGGACCACTGTTTGCCCCTCATCTCCCCACGAAGCTTCCCCCCAAG  ${\tt CCTCCAGGTAGCCAGTGGAGAGCATGGTGCTGCATTGCCGGGCACTGGCCGAACCCGAGATCTACTG}$  $\tt GTTGGCTTGTGTATGGGCCAGGACCAAAGAGGCCACTTCTTGCCACAGAGCCTTAGGGGATCGTCCTGGGCTCAT$ TGCCATCCTGGCTCTCCTCTCTCTCGCAGCTGGGCTAGCGGCCCACCTTGGCACAGGCCAACCCAGGAA GGGTGTGGGTGGGAGGCGGCCTCTCCCTCCAGCCTGGGCTTTCTGGGGGCTGGAGTGCCCCTTCTGTCCGGGTTGT GTCTGCTCCCCTCGTCCTGCCCTGGAATCCAGGGAGGAAGCTGCCCAGATCCTCAGAAGGGGAGACACTGTTGCC  ${ t ACCATTGTCTCAAAATTCT} { t TGA} { t AGCCTGTTCTCAGCAGTAGAGAAATCACTAGGACTACTTTTTACCAAA}$ AGAGAAGCAGTCTGGGCCAGATGCCCTGCCAGGAAAGGGACATGGACCCACGTGCTTGAGGCCTGGCAGCTGGGC  ${\tt CAAGACAGATGGGGCTTTGTGGGCCCTGGGGGTGCTTCTGCAGCCTTGAAAAAGTTGCCCTTACCTCCTAGGGTCA}$  ${\tt CCTCTGCTGCCATCTTCCCAAGGAACAGGAGGGACTTTGGCTAGAGCCTCCTGCCTCCCCATCTT}$  $\tt CGCCTCATCTCAGCAGCCTGGGCTCGGCATTCCGAAGCTGACTTTCTATAGGCAATTTTGTACCTTTGTGGAGAA$ 

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA VPPALPAGTQTLLLQSNSIVRVDQSELGYLANLTELDLSQNSFSDARDCDFHALPQLLSLHL EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE MLPNLEILMIGGNKVDAILDMNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELVSIDKFALVNLP ELTKLDITNNPRLSFIHPRAFHHLPQMETLMLNNNALSALHQQTVESLPNLQEVGLHGNPIR CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPFREMTDHCLPLISPRSFPPSLQ VASGESMVLHCRALAEPEPEIYWVTPAGLRLTPAHAGRRYRVYPEGTLELRRVTAEEAGLYT CVAQNLVGADTKTVSVVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNLTW SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS CHRALGDRPGLIAILALAVLLLAAGLAAHLGTGQPRKGVGGRRPLPPAWAFWGWSAPSVRVV SAPLVLPWNPGRKLPRSSEGETLLPPLSQNS

Signal sequence:

amino acids 1-18

Transmembrane domain:

amino acids 629-648

N-glycosylation site.

amino acids 94-98, 381-385, 555-559, 583-587

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 485-489

Casein kinase II phosphorylation site.

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146, 243-247, 313-317, 488-492, 700-704

Tyrosine kinase phosphorylation site.

amino acids 532-540

N-myristoylation site.

amino acids 15-21, 493-499, 566-572

Amidation site.

amino acids 470-474, 660-664, 692-696

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT CCTGGCCACACTCTTCAAGATCCTGGCGTCCTTCTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA CGACATCCCCGACGTCAAGAACGACTTCGCCTTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA GGACAAGCTCCGGCAGCGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTCATGCTCAGTGGCAT CCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCCTCAAGCTGGAGCTGATCCCGACGTGACCATCCCGCC  ${\tt CAGCATTGCCCAGCTCACGGGCCTCAAGGAGCTGTGGCTCTACCACACAGGCGGCCAAGATTGAAGCGCCTGCGCT}$ GGCCTTCCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACAACCGCTACATCGTCATCGA AGATGTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCCTCAACAGCCTCAA GAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACTCCATCTTCAGCCT  ${\tt CCACAACCTGCAGGAGATTGACCTCAAGGACAACCATCGAGGAGATCATCAGCTTCCAGCACCT}$ GCACCGCCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCAA CTACCTGGACCTCAGCCACAACAACCTGACCTTCCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCT AGCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGGCTGACAAGGAGCAGGCCTGAGCGAG GCCGGCCCAGCACAGCAGCAGGACCGCTGCCCAGTCCTCAGGCCCGGAGGGCAGGCCTAGCTTCTCCCAG AACTCCCGGACAGCCAGGACAGCCTCGCGGCTGGGCAGGAGCCTGGGGCCGCTTGTGAGTCAGGCCAGAGCGAGA GGACAGTATCTGTGGGGCTGGCCCCTTTTCTCCCTCTGAGACTCACGTCCCCCAGGGCAAGTGCTTGTGGAGGAG AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAG AGATAACTTATACATTCCCAAGAAAGTTCAGCCCAGATGGAAGGTGTTCAGGGAAAGGTGGGCTGCCTTTTCCCC CCATGGGACGGTCACCCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGCAGGCCTCCAGCTGGA AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTCCCTTGAGCAAAGCAGCCAGACGT TGAACTGTGTTTCCCTGGGCGCAGGGTGCAGGGTGTCTTCCGGATCTGGTGACCTTGGTCCAGGAGTT ATGTGTCTTGGCAGGCACTCATTTCTGTGGCTGTCGGCCAGAGGAATGTTCTGGAGCTGCCAAGGAGGAGGAG CCAGTGCCACCGCTGGCCTCCGCTGCTTCCATCAGCCCTGTCGCCACCTGGTCCTTCATGAAGAGCAGACACTTA GAGGCTGGTCGGGAATGGGGAGGTCGCCCCTGGGAGGGCAGGCGTTCCCAAGCCGGTTCCCGTCCCTGGCGC  $\tt CTGGAGTGCACACCCCAGTCGGCACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGGTCCCCACCTT$ AGAAGGGTCCCCGCCTTAGATCAATCACGTGGACACTAAGGCACGTTTTAGAGTCTCTTGTCTTAATGATTATGT  ${\tt CCATCCGTCTGTCCATTTGTGTTTTCTGCGTCGTGTCATTGGATATAATCCTCAGAAATAATGCACACTAG}$ CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAACTTGTAGACTCGGTCACAGTATCAAATAAA ATCTATAACAGAAAAAAAAAAAAA

MRQTIIKVIKFILIICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYI SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK RFAVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLELHLFMLSGIPDTVFDLVELEV LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLRENLRALHIKFTDIKEIPLWI YSLKTLEELHLTGNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIISFQ HLHRLTCLKLWYNHIAYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYLDLSHNNLT FLPADIGLLQNLQNLAITANRIETLPPELFQCRKLRALHLGNNVLQSLPSRVGELTNLTQIE LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

#### Transmembrane domain:

amino acids 51-75 (type II)

#### N-glycosylation site.

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 85-89

#### Casein kinase II phosphorylation site.

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368, 398-402, 493-497

#### N-myristoylation site.

amino acids 173-179, 261-267, 395-401, 441-447

 $\tt GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTC \underline{\textbf{ATG}} \texttt{GAGCTGGCACTGCGGCGCTCTCCCGT}$  $\tt CCCGCGGTGGTTGCTGCTGCTGCTGCTGGTCTGACTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT$ GGCCCACAGAGGAGGCCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCTGGTCAT GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTGGAAACTTTGAGGAAATTGGGC CCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT GTGGATAATCCCGTGGGCACTGGGTTCAGTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG AATTCCAGACAGTTCCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC ATTGGTCTAGAGCTTTATAAGGCCATTCAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT  ${\tt TGCCTTGGGTGATTCCTGGATCTCCCCTGTTGATTCGGTGCTCTCCTGGGGACCTTACCTGT}$ ACAGCATGTCTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT GATCATTGAACAGAACACAGATGGGGTGAACTTCTATAACATCTTAACTAAAAGCACTCCCA CGTCTACAATGGAGTCGAGTCTAGAATTCACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTCAGGAGGCCTG ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC TACTGGATTCTGAAAGCTGGTCATATGGTTCCTTCTGACCAAGGGGACATGGCTCTGAAGAT  ${\tt GATGAGACTGGTGACTCAGCAAGAA} \underline{{\tt TAG}} {\tt GATGGGGGCTGGAGATGAGCTGGTTTGGCCT}$ TGGGGCACAGAGCTGAGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT AACTGGGGCTGTGATCAAGAAGGTTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAAACCTAAGATTTTTTAAAAAATTGAT TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSDLKPRKTTWLQAASLLFVDNPVGTGFSY VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVNKGLYRE ATELWGKAEMIIEQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQRDALS QLMNGPIRKKLKIIPEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDLIVDTMGQEAWVRKLKWPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP SDQGDMALKMMRLVTQQE

### Signal sequence:

amino acids 1-25

N-glycosylation site.

amino acids 64-68, 126-130, 362-366

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 101-105

Casein kinase II phosphorylation site.

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

N-myristoylation site.

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175, 187-193, 195-201, 331-337, 332-338, 360-366

GGCCGCGGGAGAGGCC<u>ATG</u>GGCGCGCGCGGGGCGCTGCTGCTGCTGCTGCTGCTC CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACTCGGGCGTTGGCCGTGGCA GGGGAGCCTGCGCCTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG CACTCACGGCGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTG CCTTGGTGAAGCTGTCTGCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAG AGAGGATGAGGCACTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG GTTTGTGCTGGCAACGCCCAAGGCGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG GTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTTCCCTCT CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTTTGGTAATAAACACATTCCAGTTGA 

MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSLRLW
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI
YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP
SPHTLQEVQVAIINNSMCNHLFLKYSFRKDIFGDMVCAGNAQGGKDACFGDSGGPLACNKNG
LWYQIGVVSWGVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL
LGPV

### Signal sequence:

amino acids 1-18

### N-glycosylation site.

amino acids 167-171, 200-204, 273-277

### Casein kinase II phosphorylation site.

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

### N-myristoylation site.

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245, 259-265, 269-275

#### Amidation site.

amino acids 33-37

Prokaryotic membrane lipoprotein lipid attachment site. amino acids 252-263,

Serine proteases, trypsin family, histidine active site. amino acids 78-84

CTTTGCCCTCATCCTCTGGCAAATGCAGTTACAGCCCGGAGCCCGACCAGCGGAGGACGC TGCCCCCAGGCTGGGTGTCCCTGGGCCGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTCGGATCCCAG CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCCAGAAGTGCCATTCT GTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTCCC TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGGAAACCCATGTTGTAAGGTCCCCAC ATCCCTACCAGCTTCCACAGGCCTTGGCCCCCCATGTGGACTTTGTGGGGGGGACTGCACCGT TTTCCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCT GCATCTGGGGGTAACCCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGG GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCCAGTTCCTGGAGCAGTATTTCCATGAC TCAGACCTGGCTCAGTTCATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG GGACAGGAGCCCTTCCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCCTGCTCTTCGCCTCAGGTGACAGT GGGGCCGGGTGTTGGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG CCCCTATGTCACCACAGTGGGAGGCACATCCTTCCAGGAACCTTTCCTCATCACAAATGAAA TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCCACGGCCTTCATACCAGGAG GAAGCTGTAACGAAGTTCCTGAGCTCTAGCCCCCACCTGCCACCATCCAGTTACTTCAATGC CAGTGGCCGTGCCTACCCAGATGTGGCTGCACTTTCTGATGGCTACTGGGTGGTCAGCAACA GAGTGCCCATTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTTTTGGGGGGGATCCTA TCCTTGATCAATGAGCACAGGATCCTTAGTGGCCGCCCCCCTCTTGGCTTTCTCAACCCAAG GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC TGGGGAACACCAACTTCCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCCTATC AGGAGAGATGGCTTGTCCCCTGCCCTGAAGCTGGCAGTTCAGTCCCTTATTCTGCCCTGTTG GAAGCCCTGCAACCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA TGCTGTGAGCTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA TTTCACTTGATATTCACTCCCAATTCACTGCAAGGAGACCTCTACTGTCACCGTTTACTCT TTCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACTTCTCTG ACTACTCTTGTCTTCCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACA TGTAAAAAA

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS
ELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL
SIRQAELLLPGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRP
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML
LSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG
RPPLGFLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

### Signal sequence:

amino acids 1-16

### N-glycosylation site.

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

### Glycosaminoglycan attachment site.

amino acids 361-365, 408-412, 538-542

#### Casein kinase II phosphorylation site.

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

#### N-myristoylation site.

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225, 248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488, 521-527, 533-539, 549-555

 ${\tt GGGCTGCTCGGCGGAACAGTGCTCGGC}$ TTCTTTCTGCTCTGTGCTGTTGGGCAAGTGAGCCCTTACAGTGCCCCCTGGAAACCCACTTG GCCTGCATACCGCCTCCCTGTCGTCTTGCCCCAGTCTACCCTCAATTTAGCCAAGCCAGACT TTGGAGCCGAAGCCAAATTAGAAGTATCTTCTTCATGTGGACCCCAGTGTCATAAGGGAACT CCACTGCCCACTTACGAAGAGGCCAAGCAATATCTGTCTTATGAAACGCTCTATGCCAATGG CAGCCGCACAGAGACGCAGGTGGGCCATCTACATCCTCAGCAGTAGTGGAGATGGGGCCCAAC ACCGAGACTCAGGGTCTTCAGGAAAGTCTCGAAGGAAGCGGCAGATTTATGGCTATGACAGC AGGTTCAGCATTTTTGGGAAGGACTTCCTGCTCAACTACCCTTTCTCAACATCAGTGAAGTT ATCCACGGGCTGCACCGGCACCCTGGTGGCAGAGAAGCATGTCCTCACAGCTGCCACTGCA TACACGATGGAAAAACCTATGTGAAAGGAACCCAGAAGCTTCGAGTGGGCTTCCTAAAGCCC AAGTTTAAAGATGGTGGTCGAGGGGCCAACGACTCCACTTCAGCCATGCCCGAGCAGATGAA ATTTCAGTGGATCCGGGTGAAACGCACCCATGTGCCCAAGGGTTGGATCAAGGGCAATGCCA ATGACATCGGCATGGATTATGATTATGCCCTCCTGGAACTCAAAAAGCCCCCACAAGAGAAAA TTTATGAAGATTGGGGTGAGCCCTCCTGCTAAGCAGCTGCCAGGGGGCAGAATTCACTTCTC TGGTTATGACAATGACCGACCAGGCAATTTGGTGTATCGCTTCTGTGACGTCAAAGACGAGA CCTATGACTTGCTCTACCAGCAATGCGATGCCCAGCCAGGGGCCAGCGGGTCTGGGGTCTAT GTGAGGATGTGGAAGAGACAGCAGCAGAAGTGGGAGCGAAAAATTATTGGCATTTTTTCAGG GCACCAGTGGGTGGACATGAATGGTTCCCCACAGGATTTCAACGTGGCTGTCAGAATCACTC  ${f TGA}$ CACAGTGTTCCCTCCTGGCAGCAATTAAGGGTCTTCATGTTCTTATTTTAGGAGAGGCC CTTATAATCTTTTACCTATTTCTTACAATTGCAAGATGACTGGCTTTACTATTTGAAAACTG GTTTGTGTATCATATCATATCATTTAAGCAGTTTGAAGGCATACTTTTGCATAGAAATAA AAAAAATACTGATTTGGGGCAATGAGGAATATTTGACAATTAAGTTAATCTTCACGTTTTTG CAAACTTTGATTTTATTTCATCTGAACTTGTTTCAAAGATTTATATTAAATATTTGGCATA CAAGAGATATGAAAAAAAAAAAAA

MAGIPGLLFLLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIYILSSSGDGAQHRDSGSSGKS RRKRQIYGYDSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD AQPGASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW IKGNYLDCREG

#### Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 93-97, 207-211

Glycosaminoglycan attachment site.

amino acids 109-113, 316-320

Casein kinase II phosphorylation site.

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

N-myristoylation site.

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

Serine proteases, trypsin family, histidine active site.

amino acids 171-177

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTCCCCCGCCCCACCAGCCATGGTGGTTT CTGGAGCGCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC ACTGCTGCCCACTGTTTCAAGGACAACCTGAACAAACCATACCTGTTCTCTGTGCTGCTGGG CCCACCTGTGTATTCCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCGTCTCGAG CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC CCCACCTCAGACCTGCAGAAGCTGAAGGTTCCTATCATCGACTCGGAAGTCTGCAGCCAT CTGTACTGGCGGGGAGCAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT GGAGGGGGAGCGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG GCGCCTGGCTGCCGGCCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA GCTCCGCGGGCGCTCAGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG CCGCGCGCTCC<u>TAG</u>GGCGCAGCGGGACGCGGGCTCGGATCTGAAAGGCGGCCAGATCCACA TCTGGATCTGCGCCGCCCCCCGCCGTAAATAGGCTCATCTACC GGCCTCAGGCCCCCTCCAAGGCATCAGGCCCCGCCCAACGGCCTCATGTCCCCGCCCCCAC GACTTCCGGCCCCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT CTCCAAAAAAAAAA

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318</pre>

><subunit 1 of 1, 317 aa, 1 stop

><MW: 33732, pI: 7.90, NX(S/T): 1

MVVSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPPNTHCWISGWGSIQDG VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC QVDGAWLLAGIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG SGAAARS

#### Signal sequence:

amino acids 1-32

#### N-glycosylation site.

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

### Glycosaminoglycan attachment site.

amino acids 826-830

#### Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

#### Tyrosine kinase phosphorylation site.

amino acids 607-615

### N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

 ${\tt GACGGCTGGCCACC} \underline{{\tt ATG}} {\tt CACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTA}$ CTGCTGGTGGCCACCACAGGCCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGATGGT GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCGTGTGGGGCCAC AACAAGGAGCGCGGGCGCGGCGAGAATCTGTTCGCCATCACAGACGAGGGCATGGACGT GCCGCTGGCCATGGAGGGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCCAAGACAGAGAGGGATC GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT GGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGGAAACGGCCCTACCAGGAGGGGACTC CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGTGGAAACCCAGGCC CCAACTTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA AAAGTGCCCTCTAGGAGCCCAGAGACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG GGAACTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC CACACGGGGCACACCTCCTAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCCTTGCCAGGTGCAGAGGGCCCTGACA AGCCTAGCGTTGTGTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCCTCTCCTGGGA  $\tt CTACTGCTCCTCCTCTGGTGTTGGCTGGAATCTTC\underline{TGA}{A}TGGGATACCACTCAAAGGG$ TGAAGAGGTCAGCTGTCCTCCTGTCATCTTCCCCACCCTGTCCCCAGCCCCTAAACAAGATA CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCCTGAGGACTGCACACC CACTGCCTACCTGGGGCTGTCTGCCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG TGTAGCTGGGGATGGGGATTCCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTC TTTGAGTGGGGGAGGCAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT GTCCAACCTGTGAAA

MHGSCSFLMLLLPLLLLVATTGPVGALTDEEKRLMVELHNLYRAQVSPTASDMLHMRWDEE LAAFAKAYARQCVWGHNKERGRRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ MCGHYTQVVWAKTERIGCGSHFCEKLQGVEETNIELLVCNYEPPGNVKGKRPYQEGTPCSQC PSGYHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRATEASDSRKMGTPSSLATGIPAFLVTEV SGSLATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEAELPPSSEVLAS VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV SGLNSGPGHVWGPLLGLLLLPPLVLAGIF

Signal sequence:

amino acids 1-22

N-glycosylation site.

amino acids 114-118, 403-407, 409-413

Glycosaminoglycan attachment site.

amino acids 439-443

Casein kinase II phosphorylation site.

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

N-myristoylation site.

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237, 250-256

Amidation site.

amino acids 82-86, 172-176

Peroxidases proximal heme-ligand signature.

amino acids 287-298

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.

amino acids 127-138

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.

amino acids 160-172

 ${ t GTAACTGAAGTCAGGCTTTTCATTTGGGAAGCCCCCTCAACAGAATTCGGTCATTCTCCAAGTT{ t ATG}{ t GTGGACGT}$ ACTTCTGTTGTTCTCCCTCTGCTTGCTTTTCACATTAGCAGACCGGACTTAAGTCACAACAGATTATCTTTCAT CAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAATGAATTGGAGACCATTCC AAATCTGGGACCAGTCTCGGCAAATATTACACTTCTCTCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGA ACATCTGAAAGAGTTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATT TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA GCCCCAACTGCAACATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGG GCAGGAACTTCATCTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAG  ${\tt TGAGCTGGACCTAACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATAC}$  ${\tt ACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTT}$ GGATCTGAAGAACAATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAG GCGACTGATACTCCAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCA TCTAGACCTGAGTGACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATT TCAGAGCTTTGTAAATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGA TGGCTTTGTGTGTGATGATTTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTC  ${\tt CAATTTGAGTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGCTTGGAAAAAAGACAATGA}$ ACTACTGCATGATGCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCAC ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCAT CCGAGCTGGGGCCATGGCACGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAAGGA GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTCAGCAAATGC AACTCTGACTGTCCTAGAAACACCATCATTTTTGCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGC CGTCCTACAGTGCATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGT AACCGAGAGGCACTTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAA ATACACATGTGAGATGTCTAACACCCTTGGCACTGAGAGAGGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAAC CTGCGACTCCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGC  ${\tt CGTGGTTTGCTGTGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGA}$  ${\tt TTGCAGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTT}$  ${\tt AGCTGACAGGCAGGATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGG}$  ${\tt ATTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGC}$ CACAGATCTGTTCCTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGA TCCTTTTGAAACATATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTA CATAAAGAAAAAGGAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATC GTGGCCTTCACATGTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTG TCTAAACAAGTCCTCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGG TACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCC  ${\tt AAGAGCCTTTTATTTGAAAGCTCATTCTTCCCCAGACTTGGACTCTGGGTCAGAGGAAGATGGGAAAGAAGGAC}$ AGATTTTCAGGAAGAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTC GAACCAAAATTACAAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTA TTTTTTAACTTTGTTTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTT  ${\tt TTATAATGCCAGATTTCTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTT}$ TTAAATAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAA

MVDVLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSAN ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP GYFDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKM QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE FCQKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNE ISWTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ  ${\tt MKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCD}$ DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQG GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA MARLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN SAGSISANATLTVLETPSFLRPLLDRTVTKGETAVLOCIAGGSPPPKLNWTKDDSPLVVTER HFFAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPOMTAPSLDDDG WATVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLAD RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP MYLKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISW PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDA YSSFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQS YDLDT

Signal sequence:

amino acids 1-19

Transmembrane domain:

amino acids 746-765

N-glycosylation site.

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459, 628-632, 669-673, 845-849, 927-931, 939-943, 956-960

Glycosaminoglycan attachment site.

amino acids 826-830

Casein kinase II phosphorylation site.

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268, 314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653, 671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962, 1013-1017, 1019-1023, 1021-1025

Tyrosine kinase phosphorylation site.

amino acids 607-615

N-myristoylation site.

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534, 612-618, 623-629, 714-720, 873-879

GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAAGAATGAAA CCACCCCAAAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAAGTATGTTCATTTTTCTC TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT GGTGTGGTGGTGTTTTCCTTTTTTGAATTTCCCACAAGAGGAGGAAATTAATAATACATCTGCAAAGAAA TTTTATTCCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTTCTTAACCACCTGGATTTCCATCTGGATGTTGCT  $\tt GTGATCAGTCTGAAATACAACTGTTTGAATTCCAGAAGGACCAACACCAGATAAATTATGA{\color{red} ATG}{\color{blue} TTGAACAAGAT}$  $\hbox{\tt GACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAACAGGGCCCTATTTGACCCCCTGCTTGTGGTGCT}$ GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCGGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA  ${\tt CCAGTTCAGCAAGGTGATTTGTGTTCGGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT}$ GCTGAACCTCCATGAGAACCAAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAAATCCT ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGAACCTCAACACTCTGGA ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAAACTGAAGGAGCTCTGGTT GCGAAACAACCCCATTGAAAGCATCCCTTCTTATGCTTTTAACAGAATTCCTTCTTTGCGCCGACTAGACTTAGG GGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACTTGAGGTATTTGAACCTTGC CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAAACTAGATGAGCTGGATCTTTCTGGGAATCA TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACTGTGGATGATACAGTCCCAGAT TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAAC ATTACTGCCTCATGACCTCTTCACTCCCTTGCATCATCTAGAGCGGATACATTTACATCACAACCCTTGGAACTG TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTTGTGCCCGGTG TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAAATGTCGGGCCTCCACATC  ${\tt CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCGGATAGCTGT}$ GCTCAGTGATGGTACGTTAAATTTCACAAATGTAACTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA TTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCAGCAACCACTACTCCTTTCTCTTACTTTTC AACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCACAGATAACAATGTGGGTCCCACTCC AGTGGTCGACTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTCGACAGAGAAAACCTT CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTCATGAAGACTACCAAAATCATCAT TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGCACCA TCGGCAAAACCATCACGCCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC CATGGAAAGCCACCTGCCCATGCCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT CAACCACACAACAACAGTTAACACAATAAATTCAATACACAGTTCAGTGCATGAACCGTTATTGATCCGAATGAA AAAAGAAAAGAAATTTATTTATTAAAAATTCTATTGTGATCTAAAGCAGACAAAAA

MLNKMTLHPQQIMIGPRFNRALFDPLLVVLLALQLLVVAGLVRAQTCPSVCSCSNQFSKVIC VRKNLREVPDGISTNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA NLNTLELFDNRLTTIPNGAFVYLSKLKELWLRNNPIESIPSYAFNRIPSLRRLDLGELKRLS YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLDELDLSGNHLSAIRPGSFQGLMHLQKL WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLHHNPWNCNCDIL WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE LKCRASTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTVQDTGMYTCMVSNSVGN TTASATLNVTAATTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTSLTPQ STRSTEKTFTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYNSYKSPFNHTTTVNTINSIHSS VHEPLLIRMNSKDNVQETQI

### Signal sequence:

amino acids 1-44

### Transmembrane domain:

amino acids 523-543

### N-glycosylation site.

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438, 442-446, 488-492, 606-610

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 183-187

### Casein kinase II phosphorylation site.

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

### N-myristoylation site.

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243, 391-397, 422-428, 433-439, 531-537

 ${\tt AGCCGACGCTGCTCAAGCTGCTATTGCAGTTGGCAGTTCTTTTCGGTTTCCCTCCTGCTGTTTGGGGGGCA}$ GAGCGCCCATGCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGTCTT  $\tt CCCGAGCCACTCCCGTCCTGGGTCGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCC$  ${ t ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCCAAATCTGGGACCA}$ GTCTCGGCAAATATTACACTTCTCTCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG TTTCAGTCCCTTGAAACTTTGGACCTTAGCAGCAACAATATTTCAGAGCTCCAAACTGCATTTCCAGCCCTACAG CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA  $\hbox{\tt CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGGTGCTCTGAAGTCT}$  $\tt CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA$ ACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG AACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC  ${\tt AATGAAATTTCCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAAACTGAGGCGACTGATACTC}$ CAAGGAAATCGGATCCGTTCTATTACTAAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT GACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAAACTGCAACAATTGCATTTAAATACA  ${\tt AATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTTGCTGTTAGCCCAGATGGCTTTGTGTGT}$ GATGATTTCCCAAACCCCAGATCACGGTTCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTTC  ${\tt ATCTGCTCAGCTGCCAGCAGCAGTGATTCCCCAATGACTTTTGCTTGGAAAAAAGACAATGAACTACTGCATGAT}$  ${\tt GCTGAAATGGAAAATTATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCCTTCGGCTG}$  $\tt CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCATCTCCAATCACTTTGGTTCATCCTACTCTGTC$ AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTCACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCC ATGGCACGCTTGGAGTGTGCTGCTGTGGGGCACCCAGCCCCCCAGATAGCCTGGCAGAAGGATGGGGGCACAGAC TTCCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGATA  ${\tt GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTCAGCAAATGCAACTCTGACTGTC}$  $\tt CTAGAAACACCATCATTTTTGCGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGCCGTCCTACAGTGC$ ATTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAAGATGATAGCCCATTGGTGGTAACCGAGAGGCAC TTTTTTGCAGCAGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAAATACACATGTGAG  ${\tt ATGTCTAACACCCTTGGCACTGAGAGGAAACGTGCGCCTCAGTGTGATCCCCACTCCAACCTGCGACTCCCCT}$  ${\tt CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGCTGT}$ GTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACĆACACAAGGCGGAGGAATGAAGATTGCAGCATTACC AACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCATCTCAGGGAACGTTAGCTGACAGGCAG GATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACCAGTTTGTCACATCTTCAGGTGCTGGATTTTTCTTACCA  ${\tt CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTC}$  $\tt CTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGATCCTTTTGAAACA$ TATCATACAGGTTGCAGTCCTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCCAGTTACATAAAGAAAAAG GAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC TCTTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGGTACCTTTGGAAAA GCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCCAAGAGCCTTTTAT GAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACTACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC AAAAAGTTATGAAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAAACTATTTTTAACTTTG TTTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGTATTTTATAATGCCAGA TTTCTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTTGTACCATTTTTTAAATAGAAGTT 

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGGRGELGQPSGVAAERPCPTTCRCLGDLLDCSR KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY  ${\tt FDNLANTLLVLKLNRNRISAIPPKMFKLPQLQHLELNRNKIKNVDGLTFQGLGALKSLKMQR}$ NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC QKLSELDLTFNHLSRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS WTIEDMNGAFSGLDKLRRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQMK KLQQLHLNTSSLLCDCQLKWLPQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDF PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKDNELLHDAEMENYAHLRAQGGE VMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA RLECAAVGHPAPQIAWQKDGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAONSA GSISANATLTVLETPSFLRPLLDRTVTKGETAVLQCIAGGSPPPKLNWTKDDSPLVVTERHF FAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWA TVGVVIIAVVCCVVGTSLVWVVIIYHTRRRNEDCSITNTDETNLPADIPSYLSSQGTLADRQ DGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY LKGNVYGSDPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISWPS HVRKLLNTSYSHNEGPGMKNLCLNKSSLDFSANPEPASVASSNSFMGTFGKALRRPHLDAYS SFGQPSDCQPRAFYLKAHSSPDLDSGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 808-828

N-glycosylation site.

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519, 688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

Glycosaminoglycan attachment site.

amino acids 886-890

Casein kinase II phosphorylation site.

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378, 383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735, 799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022, 1073-1077, 1079-1083, 1081-1085

Tyrosine kinase phosphorylation site.

amino acids 667-675

N-myristoylation site.

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433, 513-519, 588-594, 672-678, 683-687, 774-780, 933-939

Leucine zipper pattern.

amino acids 58-80, 65-87

CAAAACTTGCGTCGCGGAGAGCCCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC TGGGGGAGCGCGTTCGGCCTGTGGGGCGCCGCTCGGCGGGGGCGCAGCAGGGGAAGCGGAAGCTGTGGTCTGCC CTGCTCCACGAGGCGCCACTGGTGTGAACCGGGAGAGCCCCTGGGTGGTCCCCTATCCCTCCTTTATATA GAAACCTTCCACACTGGGAAGGCAGCGGCGAGGCAGGAGGGCTCATGGTGAGCAAGGAGGCCGGCTGATCTGCAG  ${\tt GCGCACAGCATTCCGAGTTTACAGATTTTACAGATACCAA} \underline{{\tt ATG}} {\tt GAAGGCGAGGAGGCAGAACAGCCTGCCTGGT}$  ${ t TCCATCAGCCCTGGCGCCCCAGGCGCATCTGACTCGGCACCCCTGCAGGCACCATGGCCCAGAGCCGGGTGCTGC}$  ${\tt TGCTCCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGCCGTGAGGGCCCCAGGATTTGGCCGAAA}$ GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGCCGGTGCTGGTACTGAGCCCTGAGG AGCCCGGGCCTGGCCCAGCCGCGGTCAGCTGCCCCCGAGACTGTGCCTGTTCCCAGGAGGGCGTCGTGGACTGTG TGGAAAAGATCTACCCTGAGGAGCTCTCCCGGCTGCACCGGCTGGAGACACTGAACCTGCAAAACAACCGCCTGA CTTCCCGAGGGCTCCCAGAGAAGGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC TGACCTTGGCACCCGCTTCCTGCCAAACGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAAGATCT ATGGGCTCACCTTTGGCCAGAAGCCAAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCCGGGC TGCCGGACAACATGTTCAACGGCTCCAGCAACGTCGAGGTCCTCATCCTGTCCAGCAACTTCCTGCGCCACGTGC CCAAGCACCTGCCGCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCGGGGGCCT CCTTCTGGAAGCTCTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACCAGCTGTCTCCGGGTCCCAGCTGGGCTGC GCACCCTCATGATCCTGCACAACCAGATCACAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCCTGGAGG AGCTCAACCTCAGCTACAACCGCATCACCAGCCCACAGGTGCACCGCGACGCCTTCCGCAAGCTGCGCCTGCTGC GCTCGCTGGACCTGTCGGGCAACCGGCTGCACACGCTGCCACCTGGGCTGCCTCGAAATGTCCATGTGCTGAAGG TCAAGCGCAATGAGCTGGCTGCCTTGGCACGAGGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACCTCA CCAGCAACCGACTGCGCGAGCCCTGGGCCCCGTGCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA TCGCCGGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAACA AGATTAGTGCGGTGCCCGCCAATGCCTTCGACTCCACGCCCAACCTCAAGGGGGATCTTTCTCAGGTTTAACAAGC TGGCTGTGGGCTCCGTGGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCAACT AGGAAGAGAACAAGA<u>TAG</u>TGACAAGGTGATGCAGATGTGACCTAGGATGATGGACCGCCGGACTCTTTTCTGC TCCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCCACGGCGTGTCCCACGGCCAGACACATGC GGAACTCACAAAAGCTGGCTTTTATTCCTTTCCCATCCTATGGGGACAGGAGCCTTCAGGACTGCTGGCC  ${\tt CAGGCACTTTTCCAATGGGCAAGCCCAGTGGAGGCAGGATGGGAGAGCCCCCTGGGTGCTGCTGGGGCCTTGGGG}\\$ 

MEGEEAEQPAWFHQPWRPGASDSAPPAGTMAQSRVLLLLLLPPQLHLGPVLAVRAPGFGRS
GGHSLSPEENEFAEEEPVLVLSPEEPGPGPAAVSCPRDCACSQEGVVDCGGIDLREFPGDLP
EHTNHLSLQNNQLEKIYPEELSRLHRLETLNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLLHSN
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMILHNQITGIGREDFATTYF
LEELNLSYNRITSPQVHRDAFRKLRLLRSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRLKHLQVLDIEGNLEFGDISKD
RGRLGKEKEEEEEEEEEE

### Signal sequence:

amino acids 1-48

### N-glycosylation site.

amino acids 243-247, 310-314, 328-332, 439-443

### Casein kinase II phosphorylation site.

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

#### N-myristoylation site.

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341, 477-483, 498-502, 539-545, 548-554

### Leucine zipper pattern.

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493, 535-557

GGGAGGGGCTCCGGGCGCGCAGCAGACCTGCTCCGGCCGCGCCTCGCCGCTGTCCTCCGGGAGCGGCAG  ${\tt CAACCTGTTCCTCGCGCGCCACTGCGCTGCCCCAGGACCCGCTGCCCAAC} {\color{red} {\tt ATG}} {\scriptsize GATTTTCTCCTGGCGCTGGT}$  $\overline{\text{GCTGGTATCCTCGCTCTACCTGCAGGCGGCCGCCGAGTTCGACGGGAGGTGGCCCAGGCAAATAGTGTCATCGAT}$ TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTCGCCAGTCTTGGGGACAGTGTCAGCC TGTGTGCCAACCACGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTCATCCTGGTTATGCTGG AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA CTCCATGGCAAACTGTCAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCTGGCCT GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCATAAAGGCTTCGATCTCATGTATATTGGAGG CAAATATCAATGTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA AGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA TAATAATTGGATTCCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCCTACTCCACCACCACCACCACC CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCCAGAAAGGCCAACCACCGGACTGACAACTATAGC TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCGGCAGCCAAAGCCCCAGG GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCCTCATGCATTCAGGGGACCTGTGCCTGTCATTCAGGCA CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCCACGGAGCAGCCCTGTG GGGAAGAATGGTGGCCATGGCTGGAGGCAAACACAGATCACCTTGCGAGGGGCTGACATCAAGAGCGAATCACA AAGA<u>TGA</u>TTAAAGGGTTGGAAAAAAGATCTATGATGGAAAATTAAAGGAACTGGGATTATTGAGCCTGGAGAAG  $\overline{ ext{AGAAGACTGAGGGGCAAACCATTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG}}$ TTCTCCATATGCACTAAGAATAGAACAAGAGGAAACTGGCTTAGACTAGAGTATAAGGGAGCATTTCTTGGCAGG TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAAA AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG CATTTCCCTCTTGACCTCCTAATGGAGAGGGGATTGAAAGGGGGAAGAGCCCCACCAAATGCTGAGCTCACTGAAATA  ${f AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC}$ ATTTTCATCGGGTGCATTCTCTCTGCTTCGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC ACACCGGCAGACCTTTCCTTCACCTCATCAGTATGATTCAGTTTCTCTTATCAATTGGACTCTCCCAGGTTCCAC AGAACAGTAATATTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCATTTAACCTGGTAAAGGCAGGGCTGG  ${\tt AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTTAATGGTTCATT}$ TGATACATTGCACTAAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTTTATTATTTAATGTTTT 

MDFLLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQPRCKHGECIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGSCSSALTCSMANCQYGCDVVKGQIRCQCPSPGLHLAPDGRTCVDVDECATGRASCPRFRQCVNTFGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDGLTCVYIPKVMIEPSGPIHVPKGNGTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRPTSKPTTRPTPKPTPIPTPPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDNRVQTDPQKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVLPLGRLMHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIKSESQR

#### Signal sequence:

amino acids 1-17

#### N-glycosylation site.

amino acids 273-277

Casein kinase II phosphorylation site.

amino acids 166-170, 345-349

Tyrosine kinase phosphorylation site.

amino acids 199-206

#### N-myristoylation site.

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242, 421-427, 433-439, 462-468, 476-482

Aspartic acid and asparagine hydroxylation site.

amino acids 104-116, 186-198, 231-243

Cell attachment sequence.

amino acids 382-385

EGF-like domain cysteine pattern signature.

amino acids 75-87

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTTGCCCCCTTTAGATTGTGA  ${\tt A} {f A} {f T} {f G} {\tt T} {\tt G} {$ CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACC AACACAAGTTCACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT GATGAAGGCAATTACATCGTGAAGGTCAACATTCAGGGAAATGGAACTCTATCTGCCAGTCA GAAGATACAAGTCACGGTTGATGATCCTGTCACAAAGCCAGTGGTGCAGATTCATCCTCCCT CTGGGGCTGTGGAGTATGTGGGGAACATGACCCTGACATGCCATGTGGAAGGGGGCACTCGG CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACCAGCTCCACCTACTCCTTTTC TCCCCAAAACAATACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT GCCTGGTGAGGAACCCTGTCAGTGAAATGGAAAGTGATATCATTATGCCCCATCATATATTAT GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCCAACACCTACT CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT GCATCTGAGAAAGTAGCCCAGAAGACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC CGGCAGGCAAGATGAAACTCATTTCACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG CACAGAAAGGAAAATCATTGTCACCTTTAGCAAGTATAACTGGAATATCACTATTTTTGATT ATATCCATGTGTCTTCTTTCCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAA ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT  $\tt GTATGAAGTTATTCAGCACATCCCTGCCCAGCAGCAGCAAGACCATCCAGAG{\color{red}{\bf TGA}} ACTTTCATGG$ GCTAAACAGTACATTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT GTGCAACAGAAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG GGAGAACGAAAGTGACAGGGGTTTCCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA ATTAGTTCTACTCTACACTTTCACTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA TGTGGAAACTTTACATTGTTCGATTTTTCAGCAGACTTTGTTTTATTAAATTTTTTATTAGTG TTAAGAATGCTAAATTTATGTTTCAATTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA AGTAGTATTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA 

MWLKVFTTFLSFATGACSGLKVTVPSHTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH
TMPKYLLGSVNKSVVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQGNGTLSASQ
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT
GRQDETHFTVIITSVGLEKLAQKGKSLSPLASITGISLFLIISMCLLFLWKKYQPYKVIKQK
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRSVPASDCVSGQDLHSTV
YEVIQHIPAQQQDHPE

#### Signal sequence:

amino acids 1-18

#### Transmembrane domain:

amino acids 341-359

#### N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

#### Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

#### Tyrosine kinase phosphorylation site.

amino acids 272-280

#### N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

#### Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

GCAAGCGGCGAA<u>ATG</u>GCGCCCTCCGGGAGTCTTGCAGTTCCCCTGGCAGTCCTGGTGCTGTT GCTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTCGCGTCATCACGGACGAGA ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCT TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC TTCCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATG TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGGATCATATACT GTTTTTGCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC AGATTGCCTTTGTCCTTCAAAAAGGCGCAGACCACAGCCATACCCCTTCAAAAAAAT TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAA  ${\tt CATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCC} {\color{red}{\bf TAG}{\bf TAAATTTTATAG}}$ AACTGTGACTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCATTAAATTGAAGAGTCTA CATTCAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG GTTTAAATAGTTCTCTAATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTTGT TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTAT TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTG TTTCACTGTGTGAAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTGTCTCAAG AAATGTGTATTTCAGTGACAATTTCGTGGTCTTTTTAGAGGTATATTCCAAAATTTCCTTGT TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTC GTTTCAAACTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAAATT TTGGCCACTTTTTTCAGATTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTTT TTTCTTTTTGGATGTGAAGGTGAACATTCCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA TTTTACATTTTGAAAATTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG CATCTTCTTGTATATGTCTTAAATGTATTTTTTGTCCTCATATACAGAAAGTTCTTAATTGAT TTTACAGTCTGTAATGCTTGATGTTTTAAAATAATAACATTTTTTATATTTTTTTAAAAGACAA ACTTCATATTATCCTGTGTTCTTTCCTGACTGGTAATATTGTGTGGGATTTCACAGGTAAAA GTCAGTAGGATGGAACATTTTAGTGTATTTTTACTCCTTAAAGAGCTAGAATACATAGTTTT CACCTTAAAAGAAGGGGGAAAATCATAAATACAATGAATCAACTGACCATTACGTAGTAGAC AATTTCTGTAATGTCCCCTTCTTTCTAGGCTCTGTTGCTGTGAATCCATTAGATTTACAG TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAAATATTGTACCATT AAAGAGTTTGGATGTGTAACTTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAAACCT 

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL QPEWESFAEWGEDLEVNIAKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI NFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL ATLFSGLLLGLCMIFVADCLCPSKRRRPQPYPYPSKKLLSESAQPLKKVEEEQEADEEDVSE EEAESKEGTNKDFPQNAIRQRSLGPSLATDKS

#### Signal sequence:

amino acids 1-26

#### Transmembrane domain:

amino acids 182-201

Casein kinase II phosphorylation site.

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

Tyrosine kinase phosphorylation site.

amino acids 107-115

N-myristoylation site.

amino acids 20-26, 192-198

Amidation site.

amino acids 25-29

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTTAACTAATTCAACAAACGGGACCCTT CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGG AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAAACTCTGGGTGTCTGC ATTGCTGATGGCCTGGTTTGGTGTCCTGAGCTGTGCAGGCCGAATTCTTCACCTCTATTG GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC TAGCAAGTCAGCTGATGCTGAGGGCTACCTGGCTCACCCTGTGAATGCCTACAAACTGG TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCCTGCAGGACTCAGCTGCA GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA GAGGGGAACTTCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG AAAGCAGCTTGATGCCGGGGAGGAGGCCACCAACCAAGTCACAGGTGCTGGACTACCTCA GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGA GGAAGAGAGAAAAAACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCCAGAAGGCA TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGGATGTTTACGAGAGCCTCTGTCGTGGG GAGGGTGTCAAACTGACACCCCGTAGACAGAAGAGGGCTTTTCTGTAGGTACCACCATGGCAA TCGTCAGGTACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA CCTAAACTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTCGCCAGCTA CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAAT TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG GTGGTGCCACCGTCTTCCCTGATCTGGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT  ${\tt TGAGACCTTGTGGATCAACAGAAGTTGAC}{\tt TGA}{\tt CATCCTTTTCTGTCCTTCCCCTTCCTGGTC}$ CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCCTTTGTATGTTCCTATCAGGCT GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT GTGACTGAAGTCCCAGCCCTTCCATTCAGCCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCTTTGTACCTCAGGTGTT 

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEKELVQSLKEYILVEEAKLSKIKSWA NKMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALEDLVLQDSAAGFIANLSVQRQFFP TDEDEIGAAKALMRLQDTYRLDPGTISRGELPGTKYQAMLSVDDCFGMGRSAYNEGDYYHTV LWMEQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLLSLDPSHERAGGNLR YFEQLLEEEREKTLTNQTEAELATPEGIYERPVDYLPERDVYESLCRGEGVKLTPRRQKRLF CRYHHGNRAPQLLIAPFKEEDEWDSPHIVRYYDVMSDEEIERIKEIAKPKLARATVRDPKTG VLTVASYRVSKSSWLEEDDDPVVARVNRRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFS RRPFDSGLKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEGDYR TRHAACPVLVGCKWVSNKWFHERGQEFLRPCGSTEVD

Signal sequence:

amino acids 1-17

N-glycosylation site.

amino acids 115-119, 264-268

Glycosaminoglycan attachment site.

amino acids 490-494

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 477-481

Casein kinase II phosphorylation site.

amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270, 346-350, 365-369, 385-389, 457-461, 530-534

Tyrosine kinase phosphorylation site.

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235

AGACAGGACAATCTTCTTGGGGATGCTGGTCCTGGAAGCCAGCGGGCCTTGCTCTGTCTTTGGCCTCATTGACCC  ${\tt CCTGGGGCACCTGGCAGGGCCTACCACC} \underline{{\tt ATG}} \underline{{\tt CGACTGAGCTCCCTGTTGGCTGCGGCCAGCGCTTC}}$ CCCTCATCTTAGGGCTGTCTCTGGGGTGCAGCCTGAGCCTCCTGCGGGTTTCCTGGATCCAGGGGGAGAGAAG ATCCCTGTGTCGAGGCTGTAGGGGAGCGAGGGGCCACAGAATCCAGATTCGAGAGCTCGGCTAGACCAAAGTG ATGAAGACTTCAAACCCCGGATTGTCCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC GGTACATCCAGACAGAGCTGGGCTCCCGTGAGCGGTTGCTGGTGGCTGTCCTGACCTCCCGAGCTACACTGTCCA CTTTGGCCGTGGCTGTGAACCGTACGGTGGCCCATCACTTCCCTCGGTTACTTCACTTCACTGGGCAGCGGGGG GCCACCTTCACACACACTTTGGGGCCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGCAGGCCC  ${\tt CCCGCCTGGCAGCCCTTGCTGGCCACCTCAGCATCAACCAAGACCTGTACTTAGGCCGGGCAGAGGAGTTCATTG}$ GCGCAGGCGAGCCCGGTACTGTCATGGGGGCTTTGGCTACCTGTTGTCACGGAGTCTCCTGCTTCGTCTGC GGCCACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCCTGACGAGTGGCTTGGACGCTGCCTCATTG ACTCTCTGGGCGTCGGCTGTCTCACAGCACCAGGGGCAGCAGTATCGCTCATTTGAACTGGCCAAAAATAGGG ACCCTGAGAAGGAAGGGAGCTCGGCTTTCCTGAGTGCCTTCGCCGTGCACCCTGTCTCCGAAGGTACCCTCATGT ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGCGGGCTTACAGTGAAATAGAACAACTGCAGGCTCAGA TCCGGAACCTGACCGTGCTGACCCCGAAGGGGAGGCAGGGCTGAGCTGGCCCGTTGGGCTCCTTTCA CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACCCTTCTCCTGTGCAGATGGGG CTCCCAAGTGCCCACTACAGGGGGGTAGCAGGCGGACGTGGGTGATGCGTTGGAGACTGCCCTGGAGCAGCTCA ATCGGCGCTATCAGCCCCGCCTGCGCTTCCAGAAGCAGCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCAC GGGGCATGGAGTACACCCTGGACCTGCTGTTGGAATGTGTGACACAGCGTGGGCACCGGCGGGCCCTGGCTCGCA AGCTGGTGCTGCCACTCCTGGTGGCTGAAGCTGCTGCAGCCCCGGCTTTCCTCGAGGCGTTTGCAGCCAATGTCC TGGAGCCACGAGAACATGCATTGCTCACCCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCCGTGGAGCTCCAG ACCCATTTCTTGGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTCG CTGTGCGAGCAGAGGCCCCTTCCCAGGTGCGACTCATGGACGTGGTCTCGAAGAAGCACCCTGTGGACACTCTCT TCTTCCTTACCACCGTGTGGACAAGGCCTGGGCCCGAAGTCCTCAACCGCTGTCGCATGAATGCCATCTCTGGCT GGCAGGCCTTCTTTCCAGTCCATTTCCAGGAGTTCAATCCTGCCCTGTCACCACAGAGATCACCCCCAGGGCCCC ACCGGCAGGCTTCTGCGGAGGGCTGCTTCTACAACGCTGACTACCTGGCGGCCCGAGCCCGGCTGGCAGGTGAAC ACCTCTTTCGGGCCGTAGAGCCAGGGCTGGTGCAGAAGTTCTCCCTGCGAGACTGCAGCCCACGGCTCAGTGAAG AACTCTACCACCGCTGCCGCCTCAGCAACCTGGAGGGGCCTAGGGGGCCCGTGCCCAGCTGGCTATGGCTCTTTG CCAGGAAGGCAAGGCAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTTTAAATATGAAAATGTTATTAA ACATGTCTTCTGCC

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPDSRARLD QSDEDFKPRIVPYYRDPNKPYKKVLRTRYIQTELGSRERLLVAVLTSRATLSTLAVAVNRTV AHHFPRLLYFTGQRGARAPAGMQVVSHGDERPAWLMSETLRHLHTHFGADYDWFFIMQDDTY VQAPRLAALAGHLSINQDLYLGRAEEFIGAGEQARYCHGGFGYLLSRSLLLRLRPHLDGCRG DILSARPDEWLGRCLIDSLGVGCVSQHQGQQYRSFELAKNRDPEKEGSSAFLSAFAVHPVSE GTLMYRLHKRFSALELERAYSEIEQLQAQIRNLTVLTPEGEAGLSWPVGLPAPFTPHSRFEV LGWDYFTEQHTFSCADGAPKCPLQGASRADVGDALETALEQLNRRYQPRLRFQKQRLLNGYR RFDPARGMEYTLDLLLECVTQRGHRRALARRVSLLRPLSRVEILPMPYVTEATRVQLVLPLL VAEAAAAPAFLEAFAANVLEPREHALLTLLLVYGPREGGRGAPDPFLGVKAAAAELERRYPG TRLAWLAVRAEAPSQVRLMDVVSKKHPVDTLFFLTTVWTRPGPEVLNRCRMNAISGWQAFFP VHFQEFNPALSPQRSPPGPPGAGPDPPSPPGADPSRGAPIGGRFDRQASAEGCFYNADYLAA RARLAGELAGQEEEEALEGLEVMDVFLRFSGLHLFRAVEPGLVQKFSLRDCSPRLSEELYHR CRLSNLEGLGGRAQLAMALFEQEQANST

Signal sequence:

amino acids 1-15

Transmembrane domain:

amino acids 489-507

N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-678

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

Cell attachment sequence.

amino acids 247-250

 $\tt GTTCTAGACGCGGGAAAA{\color{red}{\textbf{ATG}}} CTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG$ AAGCATTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA TGCACCACCATGAGCATCACCACCACAAGCTCCTAACAAAGAAGATATCTTGAAAATTTCA GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAAAACC CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG AGTTCTTCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACTG GTTCTTCCTTGCACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTTGTTAA AAAAGGATCCATCACAGCCTTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT CAATATCCCAGAAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC ATCAGATGCATGTGATGTATGGGGTATACCGCCTTAGGGCATTTTGGGCATATTTTCAAT GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGGGGTAGAAAAGCG TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACTACATATCCAA TACAGCTGTATGTTTCTTTTTCTTTTTCTAATTTGGTGGCACTGGTATAACCACACATTAAAG GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT TGTGATTAAAGTAAAACTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAAATGTACAACTAAGCAACTAAAGAAA ATTAAAGTGAAAGTTGAAAAAT

MLSESSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME LSKSFRVYCIILVKPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK AYKYAFDKYRDQYNWFFLARPTTFAIIENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGG IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL PPNGSDND

#### Signal sequence:

amino acids 1-33

#### N-glycosylation site.

amino acids 121-125, 342-346

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 319-323, 464-468

Casein kinase II phosphorylation site.

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389, 399-403, 409-413, 473-477, 729-733, 748-752

Tyrosine kinase phosphorylation site.

amino acids 736-743

#### N-myristoylation site.

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550, 558-564, 651-657, 657-663, 672-672

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 14-25

#### Cell attachment sequence.

amino acids 247-250

CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTCTCCCAAATGTTCTTATGGACTGTTGCT GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTCGCAT CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTCACAGAGCTCTCCTGCT ACAATTATGGATCAGGTTCAGTCAAGAATTGTTGTCCATTGAACTGGGAATATTTTCAATCC AGCTGCTACTTCTTTCTACTGACACCATTTCCTGGGCGTTAAGTTTAAAGAACTGCTCAGC CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCCTACAAGA AACCTAAAATGAGAGAGTTTTTTATTGGACTGTCAGACCAGGTTGTCGAGGGTCAGTGGCAA TGGGTGGACGCCACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG AACAAAGGAAAATCTCTT<u>TAA</u>GAACAGAAGGCACAACTCAAATGTGTAAAGAAGGAAGAGCA AGAACATGGCCACACCCACCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC AAAAA

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN FTELSCYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQ EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCATMRDSS NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

#### Signal sequence:

amino acids 1-42

N-glycosylation site.

amino acids 2-6, 62-66, 107-111

Casein kinase II phosphorylation site.

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

N-myristoylation site.

amino acids 15-21, 74-80, 155-161

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 27-38

GGGACTACAAGCCGCGCCGCCGCCGCCGCCCCCCCAGCAACCCTCGACATGGCGCTGAGGCGGCCACCGCGAC TCCGGCTCTGCGCTGCCTGACTTCTTCCTGCTGCTGTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC TCAAATCCAGCAATCGAACCCCAGTGGTACAGGAATTTGAAAGTGTGGAACTGTCTTGCATCATTACGGATTCGC AGACAAGTGACCCCAGGATCGAGTGGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTTGACAACAAAA TTCAGGGAGACTTGGCGGGTCGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG ACTCAGCCCTTTATCGCTGTGAGGTCGTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC ACTGCCAGGAGAGTGAGGGCCACCCCCGGCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCCACGGATT CCAGAGCCAATCCCAGATTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACTTTGGTGTTCACTGCTG TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG TCACGTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA TGATC<u>TGA</u>GACCCGCGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTCAGAAGCTTTTCGTTTTGGCCAAAGTTGACCA CTACTCTTACTCTAACAAGCCACATGAATAGAAGAATTTTCCTCAAGATGGACCCGGTAAATATAACCACAA GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCTAATCTGTTTCTGGCCTGATTCCCGCATGAGTATTAGG GTGATCTTAAAGAGTTTGCTCACGTAAACGCCCGTGCTGGGCCCTGTGAAGCCAGCATGTTCACCACTGGTCGTT CAGCAGCCACGACAGCACCATGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCGGCGGAACCCA GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTTGGATCAGCATTTTGTAAAAACAACCAAAATCAGGAAG GTAAATTGGTTGCTGGAAGAGGGATCTTGCCTGAGGAACCCTGCTTGTCCAACAGGGTGTCAGGATTTAAGGAAA TACATCTAAATTTTTGCTAAGGATGTATTTTGATTATTGAAAAGAAAATTTCTATTTAAACTGTAAATATATTGT CATACAATGTTAAATAACCTATTTTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTTAAAAT TGGAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCCTTTTCT  ${\tt CACACAAGTTTTAGCCTTTTCACAAGGGAACTCATACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT}$ TAAAAATTCCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT  $\tt CCAGTCAGCTCCTGGGGTTGCGCCAGGGGCCCCCGCTCTAGCTCACTGTTGCCTCGCTGTCTGCCAGGAGGCCCT$ GCCATCCTTGGGCCCTGGCAGTGTGTCCCAGTGAGCTTTACTCACGTGGCCCTTGCTTCATCCAGCACAGC TCTCAGGTGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTTAAAAGAAAATGGAT GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTCAGGTGGAGCAGCCAGGTGAAAGGCCTGGCGGGGAGGAAAG TGAAACGCCTGAATCAAAAGCAGTTTTCTAATTTTGACTTTAAATTTTTCATCCGCCGGAGACACTGCTCCCATT TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCTGGA AATTGCATACATGAGACTGTGTTGACTTTTTTTAGTTATGTGAAACACTTTGCCGCAGGCCGCCTGGCAGAGGCA GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTGCATGCCATCCTGGATGCTTAGCATGCAAGTTC TCTTGGTTGTCATAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA TGGGAACCAGGTCTGAAAAAGTAGAGAGAGTGAAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA  $\tt CGGAAAAGGAATACTCGTGTATTTTAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT$ GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTACTAACACACCGTAATTTGGCATTTGTTTAAC CTCATTTATAAAAGCTTCAAAAAAACCCA.

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624</pre>

><subunit 1 of 1, 310 aa, 1 stop

><MW: 35020, pI: 7.90, NX(S/T): 3

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTSD PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSLKIWNVTRRDSALYRCEVVARNDRK EIDEIVIELTVQVKPVTPVCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRA NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVYDLNIGGIIGG VLVVLAVLALITLGICCAYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

Important features of the protein:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site. amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267